Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US04/015136

International filing date: 13 May 2004 (13.05.2004)

Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 60/482,992
Filing date: 27 June 2003 (27.06.2003)

Date of receipt at the International Bureau: 20 August 2004 (20.08.2004)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

1210442

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

August 11, 2004

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: ***60/482,992***
FILING DATE: ***June 27, 2003***
RELATED PCT APPLICATION NUMBER: ***PCT/US04/15136***

Certified by

Jon W Dudas

Acting Under Secretary of Commerce
for Intellectual Property
and Acting Director of the U.S.
Patent and Trademark Office

PROVISIONAL APPLICATION COVER SHEET

Sequence

This is a request for filing a PROVISIONAL APPLICATION under 37 CFR § 1.53(c).

EXPRESS MAIL LABEL NO: EV251286492US DATE OF DEPOSIT: June 27, 2003 EV251286492US	Docket Number: MON-3042	Type a plus sign (+) inside this box→	

INVENTOR(S)/APPLICANT(S)		
GIVEN NAME (FIRST AND MIDDLE [IF ANY])	**FAMILY NAME OR SURNAME**	**RESIDENCE (CITY AND EITHER STATE OR FOREIGN COUNTRY)**
Xia	LI	Havertown, Pennsylvania
Weihua	LI	Broomall, Pennsylvania
Danielle R.	REED	Glenside, Pennsylvania
Alexander Alexeyevich	BACHMANOV	Philadelphia, Pennsylvania
Joseph G.	BRAND	Wayne, Pennsylvania

☐ Additional inventors are being named on the separately numbered sheets attached hereto

TITLE OF THE INVENTION (280 characters max)

TASTE RECEPTOR OF THE T1R FAMILY FROM DOMESTIC CAT

CORRESPONDENCE ADDRESS

Attorney Name: Felicity E. Groth
WOODCOCK WASHBURN LLP
One Liberty Place - 46th Floor
Philadelphia PA 19103
Telephone (215) 568-3100
Facsimile (215) 568-3439

ENCLOSED APPLICATION PARTS (check all that apply)

☒ Specification Number of Pages: 61 ☒ Drawing(s) Number of Sheets: 13	☒ Claims (optional) ☒ Other (specify) Sequence Listing, Statement to Support

METHOD OF PAYMENT (check one)

☒ A check or money order is enclosed to cover the Provisional filing fee:
- ☐ $80.00 Small Entity
- ☒ $160.00 Large Entity

☐ The Commissioner is hereby authorized to charge filing fee and credit Deposit Account No.: 23-3050

☒ The Commissioner is hereby authorized to charge Deposit Account No. 23-3050 any fee deficiency or credit account for any overpayment.

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

☒ No.

☐ Yes, the name of the U.S. Government agency and the Government contract number are

Respectfully submitted,

SIGNATURE Felicity E. Groth

Date: June 27, 2003

TYPED or PRINTED NAME Felicity E. Groth

REGISTRATION NO. 47,042
(if appropriate)

PROVISIONAL APPLICATION FILING ONLY

TASTE RECEPTOR OF THE T1R FAMILY FROM DOMESTIC CAT

FIELD OF THE INVENTION

[0001] The present invention relates to the field of sensory mechanisms of the domestic cat, *Felis catus*. The invention relates, for example, to the discovery of a gene of *Felis catus* encoding a taste receptor of the T1R family, specifically T1R3. The invention further relates to the polypeptide encoded by the T1R3 gene of the domestic cat, and to methods and uses of the same.

BACKGROUND OF THE INVENTION

[0002] The sense of taste is important for determining food choice, for regulating food intake, and for ensuring efficient use of ingested nutrients. Taste can act as a warning system for the

presence of potentially harmful foods, by, for example, the aversive sensations of sourness or bitterness, and as an attractant to potentially nutrient-rich foods, by, for example, the appealing sensations of sweetness, saltiness, and umami.

[0003] Taste stimuli are received by taste receptor cells assembled into taste buds that are located in the epithelium of taste papillae of the tongue (Kitagawa *et al.*, *Bioch. Bioph. Res. Comm.*, 283:236-242 (2001)). The stimuli are believed to be transduced by taste receptors at the surface of the taste receptor cells (*Id.*). The taste receptors encoded by the genes of a given species are reflective of that species' food choices. For example, the "sweet receptors" of an herbivorous species are expected to be different from those of a carnivorous species, since the two consume completely different diets whose foods contain different primary stimuli. Since taste receptor specificity likely reflects food choice, it follows that receptor sequence homology among species may be as predictive or more predictive of food preferences of a given species as phylogenetic relatedness among species.

[0004] The behavior of the domestic cat (*Felis catus*), a carnivore, towards stimuli such as sweet carbohydrates, which it generally cannot taste, and towards L-amino acids, which it generally can taste, should be explicable by the specificity of taste receptors of other carnivores. Direct knowledge of taste receptor genes will allow insight into an animal's sensory world and may be useful for identifying modulators of the taste receptors encoded thereby to influence an animal's taste preferences.

[0005] Molecular receptors for the taste element of sweetness have recently been identified from human, mouse, and rat. Thus far, there are three known members of the T1R taste receptor family: T1R1, T1R2, and T1R3 (Montmayeur & Matsunami, *Curr. Opin. Neurobiol.*, 12:366-371 (2002)). The T1Rs are G-protein coupled receptors with long N-terminal extracellular domains believed to be involved in ligand binding (*Id.*). The T1R3 receptor is located within the *Sac* locus, the primary genetic locus controlling preference for sweet-tasting stimuli in mice (Li *et al.*, *Mamm. Genome*, 12:13-16 (2001); Li *et al.*, *Mamm. Genome*, 13:5-19 (2002)). Within the cell, the taste receptors heterodimerize, with T1R3 being required for the activity of T1R1 and T1R2. In mouse, the T1R1/T1R3 heterodimer functions as a receptor for selected amino acids, while the T1R2/T1R3 heterodimer functions as a receptor for stimuli considered sweet by humans. It appears that the T1R3 component couples the taste receptor to cellular signal transduction processes, thereby ensuring that the stimulus-binding event is transduced to a neural signal. Thus, knowledge of the T1R3 receptor may lead to better understanding of species-specific reactions to sapid stimuli.

[0006] Currently, mechanisms for identifying novel taste stimuli for the domestic cat are limited, for example, to exhaustive and difficult feeding studies in which a novel ingredient is paired with a control ingredient and intake of the two are compared. Considerable time, effort, and expense can be expended in the discovery of a single stimulus. Furthermore, feline illnesses often are exacerbated by a cat's refusal to eat. Additionally, the molecular features that define acceptable taste stimuli for domestic cat remain largely unknown, making rational computational design approaches for taste stimuli difficult. As a result, knowledge of the feline taste receptor and its ligands may lead to a better understanding of cat taste perception and modulation thereof.

[0007] The present invention provides, for example, a novel feline taste receptor, T1R3, and methods of use thereof to identify taste stimuli preferred by the cat. The screening methods of the invention allow the rapid screening of binding partners and/or modulators, thereby presenting significant cost savings. The results of the feline T1R3 receptor studies reflect the taste profile of the domestic cat.

SUMMARY OF THE INVENTION

[0008] Certain embodiments of the present invention relate to polynucleotides encoding a T1R3 receptor, including, but not limited to polynucleotides having the nucleotide sequence of SEQ ID NO:1, fragments of the polynucleotide of SEQ ID NO:1 encoding a polypeptide having substantially the same biological activity as a polypeptide encoded by the nucleotide sequence of SEQ ID NO:1, variants of the polynucleotide of SEQ ID NO:1 having at least 80% homology to the polynucleotide of SEQ ID NO:1 and encoding a polypeptide having substantially the same biological activity as a polypeptide encoded by the nucleotide sequence of SEQ ID NO:1, variants of the polynucleotide of SEQ ID NO:1 having at least 80% homology to the polynucleotide of SEQ ID NO:1 and encoding a polypeptide conferring modified taste perception to one or more taste stimuli relative to a polypeptide encoded by the polynucleotide of SEQ ID NO:1, nucleotide sequences encoding the amino acid sequence of SEQ ID NO:2, nucleotide sequences substantially complementary to the nucleotide sequence of SEQ ID NO:1, and nucleotide sequences that hybridize to the complement of the polynucleotide having SEQ ID NO:1 under high stringency conditions. The polynucleotides of the invention may be DNA or RNA and may be single- or double-stranded. In some embodiments of the invention, the polynucleotide fragments have at least about 42 nucleotides. The polynucleotide fragments of the invention encode, for example, an extracellular domain of the polypeptide of SEQ ID NO:2, a transmembrane domain of the polypeptide of SEQ ID NO:2, or an intracellular domain of the polypeptide of SEQ ID NO:2. In other embodiments of the invention, the polynucleotide

variants of the polynucleotide of SEQ ID NO:1 encoding an amino acid sequence of SEQ ID NO:2 having a nonconserved amino acid substitution, for example, at residue 59 and/or residue 64.

[0009] The invention also encompasses expression vectors containing the polynucleotides of the invention operably linked to a promoter. Another embodiment of the invention provides host cells containing the expression vector. The host cells may be mammalian, including feline. The invention further encompasses cell cultures of the host cells. The invention also encompasses methods of producing feline T1R3 receptor by culturing the host cells and recovering receptor therefrom.

[0010] Another embodiment of the invention includes T1R3 receptor polypeptides encoded by polynucleotides of the invention. The polypeptides of the invention include, for example, those having the amino acid sequence of SEQ ID NO:2, fragments of at least 30 contiguous amino acids of SEQ ID NO:2, and variants thereof having substantially the same biological activity as the polypeptide of SEQ ID NO:2. The variant polypeptides of the invention may have an amino acid sequence having at least one sequence variation of SEQ ID NO:2 that confers modified taste perception to one or more taste stimuli relative to a polypeptide of SEQ ID NO:2. The invention provides methods of identifying a feline T1R3 receptor variant that confers modified taste perception by expressing a variant of the polynucleotide of SEQ ID NO:1 having at least 80% homology to the polynucleotide of SEQ ID NO:1 and detecting an increase or a decrease in the biological activity of the polypeptide encoded by the variant relative to the biological activity of the polypeptide encoded by SEQ ID NO:1.

[0011] The invention further provides kits for the detection of polynucleotides encoding a feline T1R3 receptor including a polynucleotide that specifically hybridizes to a polynucleotide encoding a polypeptide having an amino acid sequence of SEQ ID NO:2 and instructions relating to detection thereof.

[0012] Also provided by the invention are antibodies that immunoreact specifically with at least one epitope of a polypeptide of the invention. The invention also includes kits for the detection of polypeptides encoding a feline T1R3 receptor including antibodies of the invention and instructions relating to detection.

[0013] Further provided by the invention are methods for identifying compounds that interact with a feline T1R3 receptor by contacting a feline T1R3 receptor with a test compound, and

detecting interaction between the receptor and the compound. The receptor may be bound to a solid support. In one aspect of the invention, the recognition sites of the receptor are coupled with a monitoring system, either electrical or optical. In another embodiment, the solid support is formulated into a feline-specific electronic tongue.

[0014] The invention also provides methods for identifying agonists and antagonists of feline T1R3 receptor. For example, the methods of the invention include identification of an agonist of a feline T1R3 receptor by expressing an expression vector of the invention in the presence of a test compound, and detecting an increase in biological activity of a polypeptide produced by the expression in the presence of the compound relative to biological activity of the polypeptide in the absence of the compound. Also included are methods for identifying agonists of a feline T1R3 receptor by contacting a polypeptide of the invention with a test compound, and detecting an increase in biological activity of the polypeptide in the presence of the compound relative to biological activity of the polypeptide in the absence of the compound.

[0015] Methods for identifying antagonists of the polypeptides of the invention also are provided. For example, the invention provides methods for identifying antagonists of a feline T1R3 receptor by expressing an expression vector in the presence of a test compound, and detecting a decrease in biological activity of a polypeptide produced by the expression step in the presence of the compound relative to biological activity of the polypeptide in the absence of the compound. Another example of methods for identifying an antagonist of a feline T1R3 receptor involves contacting a polypeptide of the invention with a test compound, and detecting a decrease in biological activity of the polypeptide in the presence of the compound relative to biological activity of the polypeptide in the absence of the compound.

[0016] Another embodiment of the invention includes compounds and compositions for modifying the taste perception of a mammal, such as a cat. The compounds and compositions may contain at least one of polynucleotides of the invention, polypeptides of the invention, or compounds identified by the methods of the invention. Examples of the compositions of the invention include veterinary foods and drinks and pharmaceutical compositions. The compositions of the invention may include a pharmaceutically acceptable excipient. The compositions of the invention may be breed-specific. Methods for modifying the taste perception of a mammal (*e.g.*, a cat) by administering to the mammal a polynucleotide of the invention, a polypeptide of the invention, and/or a compound identified according to the methods of the invention also are provided.

[0017] The invention further provides transgenic animals comprising a polynucleotide of the invention.

[0018] The materials, methods, and examples provided herein are illustrative only and are not intended to be limiting. Other features and advantages of the invention will be apparent from the following detailed description and claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0019] **Figure 1** shows the multiple sequence alignment of the T1R3 receptor of domestic cat (SEQ ID NO:1) with known nucleotide sequences of receptors of the T1R family from human (T1R1, SEQ ID NO:8; T1R2, SEQ ID NO:5; T1R3, SEQ ID NO:11), mouse (T1R1, SEQ ID NO:6; T1R2, SEQ ID NO:3; T1R3, SEQ ID NO:9), and rat (T1R1, SEQ ID NO:7; T1R2, SEQ ID NO:4; T1R3, SEQ ID NO:10). An asterisk (*) indicates a conserved nucleotide position among the sequences.

[0020] **Figure 2** shows the deduced amino acid sequence of the domestic cat taste receptor, T1R3 (SEQ ID NO:2), aligned with the amino acid sequences of members of the T1R receptor family from human (T1R1, SEQ ID NO:17; T1R2, SEQ ID NO:20; T1R3, SEQ ID NO:12), rat (T1R1, SEQ ID NO:16; T1R2, SEQ ID NO:19; T1R3, SEQ ID NO:14), and mouse (T1R1, SEQ ID NO:15; T1R2, SEQ ID NO:18; T1R3, SEQ ID NO:13). An asterisk (*) indicates a conserved nucleotide position among the sequences. A colon (:) indicates an observed conserved amino acid substitution. A period (.) indicates an observed semi-conserved amino acid substitution. The deduced amino acid sequence for cat T1R3 (SEQ ID NO:2) contains four additional amino acids at positions 11-14 relative to the homologous T1R3 receptors of mouse (SEQ ID NO:13), human (SEQ ID NO:12), and rat (SEQ ID NO:14). The deduced sequence for cat reveals a threonine in position 64, a position equivalent to amino acid 60 in mouse, and a leucine at position 59, a position equivalent to position 55 in mouse. In mouse, amino acid substitutions of a threonine at position 60 and an alanine at position 55, both positions located within the putative extracellular N-terminal domain of the polypeptide, are present in strains of mice demonstrating low preference for the sweet stimulus saccharin (Bachmanov *et al.*, *Chem. Senses*, 26:925-933 (2001)). Leucine is a conservative substitution for alanine. Accordingly, the amino acid sequence differences of cat and mouse T1R3 receptor may account for functional differences that lead to different taste preferences between the two species.

[0021] **Figure 3** illustrates a phylogenetic tree showing the relatedness of the domestic cat T1R3 receptor to the T1R family of receptors including human, rat, and mouse T1R1, T1R2, and T1R3. The T1R receptors of the rat and mouse are closely related, while the T1R3 receptor of human and cat diverge from rat and mouse. Interestingly, the sweet stimuli to which the rat and mouse respond are very similar, whereas those that stimulate the human and those that stimulate the cat differ from one another and from those for rat and mouse.

[0022] **Figure 4** illustrates the predicted conformation of cat T1R3 receptor. The cat T1R3 receptor is a seven transmembrane receptor similar in structure to other known members of the T1R family of receptors. The structure of the feline T1R3 receptor was generated through use of a protein modeling program available at <www.ebi.ac.uk/~moeller/transmembrane.html>.

DETAILED DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

[0023] The reference works, patents, patent applications, and scientific literature that are referred to herein reflect in part the knowledge of those with skill in the art and are hereby incorporated by reference in their entirety to the same extent as if each was specifically and individually indicated to be incorporated by reference. Any conflict between any reference cited herein and the specific teachings of this specification shall be resolved in favor of the latter. Likewise, any conflict between an art-understood definition of a word or phrase and a definition of the word or phrase as specifically taught in this specification shall be resolved in favor of the latter.

[0024] Standard reference works setting forth the general principles of recombinant DNA technology are known to those of skill in the art (Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, 1998; Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, 2D ED., Cold Spring Harbor Laboratory Press, Plainview, New York, 1989; Kaufman et al., Eds., HANDBOOK OF MOLECULAR AND CELLULAR METHODS IN BIOLOGY AND MEDICINE, CRC Press, Boca Raton, 1995; McPherson, Ed., DIRECTED MUTAGENESIS: A PRACTICAL APPROACH, IRL Press, Oxford, 1991).

[0025] As used herein, "taste perception" refers to a response (*e.g.*, biochemical, behavioral) or sensitivity of a T1R3 receptor of the invention to a taste stimulus. "Taste stimulus" as used herein refers to any compound that elicits, for example at the biochemical level (*e.g.*, activation or inhibition of a taste receptor) or behavioral level, a taste response which would be perceived by a mammal as at least one of the five taste elements, including sweet, salty, sour, bitter, and

umami. "Taste perception" or "taste stimulus," or variants thereof, does not require, though it does include, transmission of a neural signal resulting in *in vivo* sensation of taste by a mammal. Modification of taste perception includes an alteration of (enhancement of, reduction to, or change to) a biochemical response, an ingestive response, a taste preference, or general behavior of a mammal in response to a compound.

[0026] As used herein "polynucleotide" refers to a nucleic acid molecule and includes genomic DNA, cDNA, RNA, mRNA, mixed polymers, recombinant nucleic acids, fragments and variants thereof, and the like. Polynucleotide fragments of the invention comprise at least 10, and preferably at least 12, 14, 16, 18, 20, 25, 30, 35, 40, 45, 50, 75, or 100 consecutive nucleotides of a reference polynucleotide. The polynucleotides of the invention include sense and antisense strands. The polynucleotides of the invention may be naturally occurring or non-naturally occurring polynucleotides. A "synthesized polynucleotide" as used herein refers to polynucleotides produced by purely chemical, as opposed to enzymatic, methods. "Wholly" synthesized DNA sequences are therefore produced entirely by chemical means, and "partially" synthesized DNAs embrace those wherein only portions of the resulting DNA were produced by chemical means. The polynucleotides of the invention may be single- or double-stranded. The polynucleotides of the invention may be chemically modified and may contain non-natural or derivatized nucleotide bases as will be readily appreciated by those skilled in the art. Such modifications include, for example, labels, methylation, substitution of one or more nucleotides with an analog, internucleotide modifications such as uncharged linkages (*e.g.*, methyl phosphonates, phosphotriesters, phosphoramidates, carbamates, *etc.*), charged linkages (*e.g.*, phosphorothioates, phosphorodithioates, *etc.*), pendent moieties (*e.g.*, polypeptides, *etc.*), intercalators (*e.g.*, acridine, psoralen, *etc.*), chelators, alkylators, and modified linkages (*e.g.*, alpha anomeric nucleic acids, *etc.*). Also included are synthetic molecules that mimic polynucleotides in their ability to bind to a designated sequence via hydrogen bonding and other chemical interactions. Such molecules are known in the art and include, for example, those in which peptide linkages substitute for phosphate linkages in the backbone of the molecule.

[0027] "Recombinant nucleic acid" is a nucleic acid generated by combination of two segments of nucleotide sequence. The combination may be, for example, by chemical means or by genetic engineering.

[0028] As used herein, "polynucleotide amplification" refers to a broad range of techniques for increasing the number of copies of specific polynucleotide sequences. Typically, amplification of either or both strand(s) of the target nucleic acid comprises the use of one or more nucleic

acid-modifying enzymes, such as a DNA polymerase, ligase, RNA polymerase, or RNA-dependent reverse transcriptase. Examples of polynucleotide amplification include, but are not limited to, polymerase chain reaction (PCR), nucleic acid sequence based amplification (NASB), self-sustained sequence replication (3SR), strand displacement activation (SDA), ligase chain reaction, Qβ replicase system, and the like. A wide variety of alternative cloning and *in vitro* amplification methodologies are well known to those skilled in the art. Examples of these techniques are found in, for example, Berger *et al.*, *Guide to Molecular Cloning Techniques*, METHODS IN ENZYMOLOGY 152, Academic Press, Inc., San Diego, CA (Berger), which is incorporated herein by reference in its entirety.

[0029] As used herein, the term "oligonucleotide" or "primer" refers to a series of linked nucleotide residues which has a sufficient number of bases to be used in a polymerase chain reaction (PCR). This short sequence is based on (or designed from) a genomic or cDNA sequence and is used to amplify, confirm, or reveal the presence of an identical, similar, or complementary DNA or RNA in a particular cell or tissue. Oligonucleotides comprise portions of a nucleic acid sequence having at least about 10 nucleotides and as many as about 50 nucleotides, often about 12 or 15 to about 30 nucleotides. They are chemically synthesized and may be used as probes. "Primer pair" refers to a set of primers including a 5' upstream primer that hybridizes with the 5' end of a target sequence to be amplified and a 3' downstream primer that hybridizes with the complement of the 3' end of the target sequence to be amplified.

[0030] As used herein, the term "probe" refers to nucleic acid sequences of variable length, for example between at least about 10 and as many as about 6,000 nucleotides, depending on use. Probes are used in the detection of identical, similar, or complementary target nucleic acid sequences, which target sequences may be single- or double-stranded. Longer probes are usually obtained from a natural or recombinant source, are highly specific, and are much slower to hybridize than oligomers, or shorter probes. They may be single- or double-stranded and are carefully designed to have specificity in PCR, hybridization membrane-based, or ELISA-like technologies. An "overgo probe" is a DNA probe comprising two short, overlapping DNA sequences (*e.g.*, 10-50 nucleotides each) with a complementary overlapping region (*e.g.*, 5-15 nucleotides) that is used in an overgo hybridization strategy. For example, an overgo probe may be two 22mers with an 8 bp complementary overlap, resulting in a 36mer overgo probe. As another example, an overgo probe may be two 24mers with an 8 bp complementary overlap, resulting in a 40mer overgo probe.

[0031] As used herein, the phrase "stringent hybridization conditions" or "stringent conditions" refers to conditions under which a probe, primer, or oligonucleotide will hybridize to its target sequence, but to a minimal number of other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences will hybridize with specificity to their proper complements at higher temperatures. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. Since the target sequences are generally present in excess, at T_m, 50% of the probes are hybridized to their complements at equilibrium. Stringent temperature conditions will generally include temperatures in excess of 30°C, typically in excess of 37°C, and may be in excess of 45°C. Stringent salt conditions will ordinarily be less than 1.0 M, typically less than 0.5 M, and may be less than 0.2 M. Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes, primers, or oligonucleotides (*e.g.*, 10 to 50 nucleotides) and at least about 60°C for longer probes, primers, or oligonucleotides. Stringent conditions may also be achieved with the addition of destabilizing agents, such as formamide.

[0032] As used herein "antisense oligonucleotide" refers to a nucleic acid molecule that is complementary to at least a portion of a target nucleotide sequence of interest and specifically hybridizes to the target nucleotide sequence under physiological conditions. The term "double stranded RNA" or "dsRNA" as used herein refers to a double-stranded RNA molecule capable of RNA interference, including short interfering RNA (siRNA) (see for example, Bass, *Nature*, 411, 428-429 (2001); Elbashir *et al.*, *Nature*, 411, 494-498 (2001)).

[0033] As used herein, the term "complementary" refers to Watson-Crick basepairing between nucleotide units of a nucleic acid molecule.

[0034] The term "marker gene" or "reporter gene" refers to a gene encoding a product that, when expressed, confers a phenotype at the physical, morphologic, or biochemical level on a transformed cell that is easily identifiable, either directly or indirectly, by standard techniques and includes, but is not limited to, genes encoding proteins that confer resistance to toxins or antibiotics such as ampicillin, neomycin, and methotroxate; genes encoding proteins that complement auxotrophic deficiencies; and genes encoding proteins that supply critical components not available from complex media. Examples of marker genes include green

fluorescent protein (GFP), red fluorescent protein (DsRed), alkaline phosphatase (AP), β-lactamase, chloramphenicol acetyltransferase (CAT), adenosine deaminase (ADA), aminoglycoside phosphotransferase (neor, G418r) dihydrofolate reductase (DHFR), hygromycin-B-phosphotransferase (HPH), thymidine kinase (TK), lacZ (encoding β-galactosidase), luciferase (luc), and xanthine guanine phosphoribosyltransferase (XGPRT). As with many of the standard procedures associated with the practice of the invention, skilled artisans will be aware of additional sequences that can serve the function of a marker or reporter. Thus, this list is merely meant to show examples of what can be used and is not meant to limit the invention.

[0035] As used herein, the term "promoter" refers to a regulatory element that regulates, controls, or drives expression of a nucleic acid molecule of interest and can be derived from sources such as from adenovirus, SV40, parvoviruses, vaccinia virus, cytomegalovirus, or mammalian genomic DNA. Examples of suitable promoters include, but are not limited to, CMV, MSH2, trp, lac, phage, and TRNA promoters. Suitable promoters that can be used in yeast include, but are not limited to, such constitutive promoters as 3-phosphoglycerate kinase and various other glycolytic enzyme gene promoters such as enolase or glyceraldehydes-3-phosphate dehydrogenase, or such inducible promoters as the alcohol dehydrogenase 2 promoter or metallothionine promoter. Again, as with many of the standard procedures associated with the practice of the invention, skilled artisans will be aware of additional promoters that can serve the function of directing the expression of a marker or reporter. Thus, the list is merely meant to show examples of what can be used and is not meant to limit the invention.

[0036] "Operably linked" refers to juxtaposition wherein the components are in a functional relationship. For example, a promoter is operably linked or connected to a coding sequence if it controls the transcription or expression of the sequence.

[0037] The terms "polypeptide," "peptide," and "protein" are used interchangeably herein. "Polypeptide" refers to a polymer of amino acids without referring to a specific length. Polypeptides of the invention include peptide fragments, derivatives, and fusion proteins. Peptide fragments preferably have at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, or 100 amino acids. Some peptide fragments of the invention are biologically active. Biological activities include immunogenicity, ligand binding, and activity associated with the reference peptide. Immunogenic peptides and fragments of the invention generate an epitope-specific immune response, wherein "epitope" refers to an immunogenic determinant of a peptide and preferably contains at least three, five, eight, nine, ten, fifteen, twenty, thirty, forty, forty-five, or fifty amino acids. Some immunogenic peptides of the invention generate an immune response

specific to that peptide. Polypeptides of the invention include naturally occurring and non-naturally occurring peptides. The term includes modified polypeptides (wherein examples of such modifications include glycosylation, acetylation, phosphorylation, carboxylation, ubiquitination, labeling, *etc.*), analogs (such as non-naturally occurring amino acids, substituted linkages, *etc.*), and functional mimetics. A variety of methods for labeling polypeptides are well known in the art and include radioactive isotopes such as ^{32}P or ^{35}S, ligands that bind to labeled antiligands (*e.g.*, antibodies), fluorophores, chemiluminescent agents, enzymes, and antligands.

[0038] As used herein, the term "amino acid" denotes a molecule containing both an amino group and a carboxyl group. In some embodiments, the amino acids are α-, β-, γ- or δ-amino acids, including their stereoisomers and racemates. As used herein the term "L-amino acid" denotes an α-amino acid having the L configuration around the α-carbon, that is, a carboxylic acid of general formula CH(COOH)(NH2)-(side chain), having the L-configuration. The term "D-amino acid" similarly denotes a carboxylic acid of general formula CH(COOH)(NH2)-(side chain), having the D-configuration around the α-carbon. Side chains of L-amino acids include naturally occurring and non-naturally occurring moieties. Non-naturally occurring (*i.e.*, unnatural) amino acid side chains are moieties that are used in place of naturally occurring amino acid side chains in, for example, amino acid analogs. Amino acid substituents may be attached, for example, through their carbonyl groups through the oxygen or carbonyl carbon thereof, or through their amino groups, or through functionalities residing on their side chain portions.

[0039] The amino acid sequences are presented in the amino (N) to carboxy (C) direction, from left to right. The N-terminal α-amino group and the C-terminal β-carboxy groups are not depicted in the sequence. The nucleotide sequences are presented by single strands only, in the 5' to 3' direction, from left to right. Nucleotides and amino acids are represented in the manner recommended by the IUPAC-IUB Biochemical Nomenclature Commission, or amino acids are represented by their three letters code designations.

[0040] As used herein, the term "antibody" is meant to refer to complete, intact antibodies, and Fab, Fab', $F(ab)_2$, F_v, and other fragments thereof. Complete, intact antibodies include antibodies such as polyclonal antibodies, monoclonal antibodies, chimeric antibodies, and humanized antibodies, felinized antibodies, and immunologic binding equivalents thereof. The antibodies of the invention may be labeled or unlabeled. Examples of labels of antibodies include, but are not limited to, radionuclides, enzymes, substrates, cofactors, inhibitors,

fluorescent agents, chemiluminescent agents, magnetic particles, and the like. Recombinant immunoglobulins are included in the invention.

[0041] As used herein, the term "binding" means the physical or chemical interaction between two proteins or compounds or associated proteins or compounds or combinations thereof. Binding includes ionic, non-ionic, Hydrogen bonds, Van der Waals, hydrophobic interactions, *etc.* The physical interaction, the binding, can be either direct or indirect, indirect being through or due to the effects of another protein or compound. Direct binding refers to interactions that do not take place through or due to the effect of another protein or compound but instead are without other substantial chemical intermediates. Binding may be detected in many different manners. As a non-limiting example, the physical binding interaction between two molecules can be detected using a labeled compound. Other methods of detecting binding are well-known to those of skill in the art.

[0042] As used herein, the term "contacting" means bringing together, either directly or indirectly, a compound into physical proximity to a molecule of interest. Contacting may occur, for example, in any number of buffers, salts, solutions, or in a cell or cell extract.

[0043] As used herein, the terms "modulates" or "modifies" means an increase or decrease in the amount, quality, or effect of a particular activity or protein. "Modulators" refer to any inhibitory or activating molecules identified using *in vitro* and *in vivo* assays for, *e.g.*, agonists, antagonists, and their homologs, including fragments, variants, and mimetics, as defined herein, that exert substantially the same biological activity as the molecule. "Inhibitors" or "antagonists" are modulating compounds that reduce, decrease, block, prevent, delay activation, inactivate, desensitize, or downregulate the biological activity or expression of a molecule or pathway of interest. "Inducers," "activators," or "agonists" are modulating compounds that increase, induce, stimulate, open, activate, facilitate, enhance activation, sensitize, or upregulate a molecule or pathway of interest. In some preferred embodiments of the invention, the level of inhibition or upregulation of the expression or biological activity of a molecule or pathway of interest refers to a decrease (inhibition or downregulation) or increase (upregulation) of greater than about 50%, 60%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%. The inhibition or upregulation may be direct, *i.e.*, operate on the molecule or pathway of interest itself, or indirect, *i.e.*, operate on a molecule or pathway that affects the molecule or pathway of interest.

[0044] A “substantially pure” polynucleotide or polypeptide is substantially separated from other cellular components that naturally accompany a native (or wild-type) nucleic acid or polypeptide and/or from other impurities (*e.g.*, agarose gel). A substantially pure protein will comprise about 60% to more than 99% w/w of a protein sample, and may be about 90%, about 95%, or about 98% pure. As used herein, the term "isolated" refers to a molecule that has been removed from its native environment. Examples of isolated nucleic acid molecules include, but are not limited to, recombinant DNA molecules contained in a vector, recombinant DNA molecules maintained in a heterologous host cell, partially or substantially purified nucleic acid molecules, and synthetic DNA or RNA molecules.

[0045] “About” as used herein refers to +/- 10% of the reference value.

[0046] As used herein, “variant” nucleotide or amino acid sequences refer to homologs, including, for example, isoforms, species variants, allelic variants, and fragments of the sequence of interest. "Homologous nucleotide sequence" or "homologous amino acid sequence," or variations thereof, refers to sequences characterized by a homology, at the nucleotide level or amino acid level, of at least about 60%, at least about 70%, at least about 75%, at least about 80%, at least about 81%, at least about 82%, at least about 83%, at least about 84%, at least about 85%, preferably at least about 90%, at least about 95%, at least about 98%, or at least about 99%, and more preferably 100%, to a reference sequence, or portion or fragment thereof encoding or having a functional domain, including, for example, but not limited to the nucleic acid sequence of SEQ ID NO:1, or a portion of SEQ ID NO:1, which encodes a functional domain of the encoded polypeptide, SEQ ID NO:2, or to the polypeptide having amino acid sequence SEQ ID NO:2, or fragments thereof having functional domains of the full-length polypeptide. Examples of functional domains of the T1R3 polypeptide of SEQ ID NO:2 include the extracellular domains (residues 1-571, 628-641, 705-730, and 787-794 of SEQ ID NO:2), the transmembrane domains (residues 572-594, 610-627, 642-664, 681-704, 731-754, 767-780, and 795-812 of SEQ ID NO:2), and the intracellular domains (residues 595-609, 665-680, 755-766, and 813-865 of SEQ ID NO:2). Isoforms can be expressed in different tissues of the same organism as a result of, for example, alternative splicing of RNA. Alternatively, isoforms can be encoded by different genes. Homologous nucleotide sequences include nucleotide sequences encoding for a species variant of a protein. Homologous nucleotide sequences also include, but are not limited to, naturally occurring allelic variations and mutations of the nucleotide sequences set forth herein. Study of mutations and polymorphisms of the T1R3 receptor polynucleotide sequence may explain breed-specific and/or individual taste preferences of a

mammal such as a cat. Additionally, sequence variants of the T1R3 receptor may be associated with specific disease states, such that knowledge of the gene allows diagnosis and treatment of T1R3-associated disorders (*e.g.*, obesity, diabetes). Homologous amino acid sequences include those amino acid sequences which encode conservative amino acid substitutions in polypeptides having amino acid sequence of SEQ ID NO:2, as well as in polypeptides identified according to the methods of the invention. Percent homology may be determined by, for example, the Gap program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, Madison Wis.), using the default settings, which uses the algorithm of Smith and Waterman (Smith and Waterman, *Adv. Appl. Math.*, 2: 482-489, 1981). Nucleic acid fragments of the invention preferably have at least about 5, at least about 10, at least about 15, at least about 20, at least about 25, at least about 50, or at least about 100 nucleotides of the reference nucleotide sequence. The nucleic acid fragments of the invention may encode a polypeptide having at least one biological property, or function, that is substantially similar to a biological property of the polypeptide encoded by the full-length nucleic acid sequence.

[0047] As is well known in the art, because of the degeneracy of the genetic code, there are numerous DNA and RNA molecules that can code for the same polypeptide as that encoded by a nucleotide sequence of interest. The present invention, therefore, contemplates those other DNA and RNA molecules which, on expression, encode a polypeptide encoded by the nucleic acid molecule of interest. DNA and RNA molecules other than those specifically disclosed herein characterized simply by a change in a codon for a particular amino acid, are within the scope of this invention.

[0048] Amino acid "insertions", "substitutions" or "deletions" are changes to or within an amino acid sequence. The variation allowed in a particular amino acid sequence may be experimentally determined by producing the peptide synthetically or by systematically making insertions, deletions, or substitutions of nucleotides in the nucleic acid sequence using recombinant DNA techniques. Alterations of the naturally occurring amino acid sequence can be accomplished by any of a number of known techniques. For example, mutations can be introduced into the polynucleotide encoding a polypeptide at particular locations by procedures well known to the skilled artisan, such as oligonucleotide-directed mutagenesis, which is described by U.S. Pat. Nos. 4,518,584 and 4,737,462.

[0049] A polypeptide variant of the present invention may exhibit substantially the biological activity of a naturally occurring reference polypeptide. "Biological activity" as used herein refers to the level of a particular function (for example, enzymatic activity) of a molecule or pathway of

interest in a biological system. "Wild-type biological activity" refers to the normal level of function of a molecule or pathway of interest. "Reduced biological activity" refers to a decreased level of function of a molecule or pathway of interest relative to a reference level of biological activity of that molecule or pathway. For example, reduced biological activity may refer to a decreased level of biological activity relative to the wild-type biological activity of a molecule or pathway of interest. "Increased biological activity" refers to an increased level of function of a molecule or pathway of interest relative to a reference level of biological activity of that molecule or pathway. For example, increased biological activity may refer to an increased level of biological activity relative to the wild-type biological activity of a molecule or pathway of interest. Reference to exhibiting "substantially the biological activity of a naturally occurring polypeptide" indicates that variants within the scope of the invention can comprise conservatively substituted sequences, meaning that one or more amino acid residues of a polypeptide are replaced by different residues that do not alter the secondary and/or tertiary structure of the polypeptide. Such substitutions may include the replacement of an amino acid by a residue having similar physicochemical properties, such as substituting one aliphatic residue (Ile, Val, Leu or Ala) for another, or substitution between basic residues Lys and Arg, acidic residues Glu and Asp, amide residues Gln and Asn, hydroxyl residues Ser and Tyr, or aromatic residues Phe and Tyr. Further information regarding making phenotypically silent amino acid exchanges are known in the art (Bowie *et al.*, *Science*, 247: 1306-1310, 1990). Other polypeptide homologs which might retain substantially the biological activities of the reference polypeptide are those where amino acid substitutions have been made in areas outside functional regions of the protein.

[0050] A nucleotide and/or amino acid sequence of a nucleic acid molecule or polypeptide employed in the invention or of a compound identified by the screening method of the invention may be used to search a nucleotide and amino acid sequence databank for regions of similarity using Gapped BLAST (Altschul *et al.*, *Nuc. Acids Res.*, 25: 3389, 1997). Briefly, the BLAST algorithm, which stands for Basic Local Alignment Search Tool is suitable for determining sequence similarity (Altschul *et al.*, *J Mol. Biol.*, 215: 403-410, 1990). Software or performing BLAST analyses is publicly available through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pair (HSPs) by identifying short words of length W in the query sequence that either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, *J Mol. Biol.*, 215: 403-410, 1990). These initial neighborhood word hits act as

seeds for initiating searches to find HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Extension for the word hits in each direction are halted when: 1) the cumulative alignment score falls off by the quantity X from its maximum achieved value; 2) the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or 3) the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLAST program uses as defaults a word length (W) of 11, the BLOSUM62 scoring matrix (Henikoff *et al.*, *Proc. Natl. Acad. Sci. USA*, 89: 10915-10919, 1992) alignments (B) of 50, expectation (E) of 10, M=5, N=4, and a comparison of both strands. The BLAST algorithm (Karlin *et al.*, *Proc. Natl. Acad. Sci. USA*, 90: 5873-5787, 1993) and Gapped BLAST perform a statistical analysis of the similarity between two sequences. One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a gene or cDNA if the smallest sum probability in comparison of the test nucleic acid to the reference nucleic acid is less than about 1, preferably less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

[0051] The term "mimetic" as used herein refers to a compound that is sterically similar to a reference compound. Mimetics are structural and functional equivalents to the reference compounds.

[0052] The terms "patient" and "subject" are used interchangeably herein and include, but are not limited to, avians, felines, canines, bovines, ovines, porcines, equines, rodents, simians, and humans. "Host cell" includes, for example, a mammalian cell (*e.g.*, human, rodent, feline), yeast cell, or plant cell. "Rodents" include, for example, rats and mice.

[0053] The term "treatment" as used herein refers to any indicia of success of prevention, treatment, or amelioration of a disease or condition. Treatment includes any objective or subjective parameter, such as, but not limited to, abatement, remission, normalization of receptor activity, reduction in the number of infectious particles in a patient, reduction in the number or severity of symptoms or side effects, an increase in the tolerance of the patient to an infection, or slowing of the rate of degeneration or decline of the patient. Treatment also includes a prevention of the onset of symptoms in a patient that may be at increased risk of infection but does not yet experience or exhibit symptoms thereof.

[0054] As used herein, the term "compound" means any identifiable chemical or molecule, including, but not limited to a small molecule, peptide, protein, sugar, nucleotide, or nucleic acid. Such compound can be natural or synthetic.

Polynucleotides

[0055] The invention provides purified and isolated polynucleotides (*e.g.*, cDNA, genomic DNA, synthetic DNA, RNA, or combinations thereof, whether single- or double-stranded) that comprise a nucleotide sequence encoding the amino acid sequence of the polypeptides of the invention. Such polynucleotides are useful for recombinantly expressing the receptor and also for detecting expression of the receptor in cells (*e.g.*, using Northern hybridization and *in situ* hybridization assays). Such polynucleotides also are useful in the design of antisense and other molecules for the suppression of the expression of T1R3 receptor in a cultured cell, a tissue, or an animal; for therapeutic purposes; or to provide a model for diseases or conditions characterized by aberrant T1R3 expression. Specifically excluded from the definition of polynucleotides of the invention are entire isolated, non-recombinant native chromosomes of host cells. One embodiment of polynucleotides of the invention is the nucleotide sequence of SEQ ID NO:1. It will be appreciated that numerous other polynucleotide sequences exist that also encode the T1R3 receptor of the invention due to the well-known degeneracy of the universal genetic code.

[0056] The invention also provides a purified and isolated polynucleotide comprising a nucleotide sequence that encodes a mammalian (*e.g.*, feline) polypeptide, wherein the polynucleotide hybridizes to a polynucleotide having a sequence of SEQ ID NO:1, or the non-coding strand complementary thereto, under stringent hybridization conditions.

[0057] Genomic DNA of the invention comprises the protein-coding region for a polypeptide of the invention and is also intended to include allelic variants thereof. It is widely understood that, for many genes, genomic DNA is transcribed into RNA transcripts that undergo one or more splicing events wherein intron (*i.e.*, non-coding regions) of the transcripts are removed, or "spliced out." RNA transcripts that can be spliced by alternative mechanisms, and therefore be subject to removal of different RNA sequences but still encode a T1R3 polypeptide, are referred to in the art as splice variants which are embraced by the invention. Splice variants comprehended by the invention therefore are encoded by the same original genomic DNA sequences but arise from distinct mRNA transcripts. Allelic variants are modified forms of a wild-type gene sequence, the modification resulting from recombination during chromosomal

segregation or exposure to conditions which give rise to genetic mutation. Allelic variants, like wild type genes, are naturally occurring sequences (as opposed to non-naturally occurring variants that arise from *in vitro* manipulation).

[0058] The invention also comprehends cDNA that is obtained through reverse transcription of an RNA polynucleotide encoding a T1R3 receptor (conventionally followed by second strand synthesis of a complementary strand to provide a double-stranded DNA).

[0059] One embodiment of the DNA of the invention comprises a double-stranded molecule along with the complementary molecule (the "non-coding strand" or "complement") having a sequence unambiguously deducible from the coding strand according to Watson-Crick base-pairing rules for DNA.

[0060] The present invention includes fragments of nucleotide sequences encoding a T1R3 receptor comprising at least 10, and preferably at least 12, 14, 16, 18, 20, 25, 30, 35, 40, 45, 50, 75, or 100 consecutive nucleotides of a polynucleotide encoding T1R3. Fragment polynucleotides of the invention may comprise sequences unique to the T1R3-encoding polynucleotide sequence, and therefore hybridize under highly stringent or moderately stringent conditions only (*i.e.*, "specifically") to polynucleotides encoding T1R3 (or fragments thereof). Polynucleotide fragments of genomic sequences of the invention comprise not only sequences unique to the coding region, but also include fragments of the full-length sequence derived from introns, regulatory regions, and/or other non-translated sequences. Sequences unique to polynucleotides of the invention are recognizable through sequence comparison to other known polynucleotides, and can be identified through use of alignment programs routinely utilized in the art, *e.g.*, those made available in public sequence databases. Such sequences also are recognizable from Southern hybridization analyses to determine the number of fragments of genomic DNA to which a polynucleotide will hybridize. Polynucleotides of the invention can be labeled in a manner that permits their detection, including radioactive, fluorescent, and enzymatic labeling.

[0061] Fragment polynucleotides are particularly useful as probes for detection of full-length or fragments of T1R3 polynucleotides. One or more polynucleotides can be included in kits that are used to detect the presence of a polynucleotide encoding T1R3, or used to detect variations in a polynucleotide sequence encoding T1R3.

[0062] The invention also embraces DNAs encoding T1R3 polypeptides that hybridize under moderately stringent or high stringency conditions to the non-coding strand, or complement, of the polynucleotides.

[0063] Exemplary highly stringent hybridization conditions are as follows: hybridization at 42°C in a hybridization solution comprising 50% formamide, 1% SDS, 1 M NaCl, 10% Dextran sulfate, and washing twice for 30 minutes at 60°C in a wash solution comprising 0.1 X SSC and 1% SDS. It is understood in the art that conditions of equivalent stringency can be achieved through variation of temperature and buffer, or salt concentration as described, for example, in Ausubel *et al.* (Eds.), PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons (1994), pp. 6.0.3 to 6.4.10. Modifications in hybridization conditions can be empirically determined or precisely calculated based on the length and the percentage of guanosine/cytosine (GC) base pairing of the probe. The hybridization conditions can be calculated as described, for example, in Sambrook *et al.*, (Eds.), MOLECULAR CLONING: A LABORATORY_MANUAL, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York (1989), pp. 9.47 to 9.51.

[0064] With the knowledge of the nucleotide sequence information disclosed in the present invention, one skilled in the art can identify and obtain nucleotide sequences which encode T1R3 from different sources (*i.e.*, different tissues or different organisms) through a variety of means well known to the skilled artisan and as disclosed by, for example, Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, Second Edition, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989).

[0065] For example, DNA that encodes T1R3 may be obtained by screening mRNA, cDNA, or genomic DNA with oligonucleotide probes generated from the T1R3 gene sequence information provided herein. Probes may be labeled with a detectable group, such as a fluorescent group, a radioactive atom or a chemiluminescent group in accordance with procedures known to the skilled artisan and used in conventional hybridization assays, as described by, for example, Sambrook *et al.*

[0066] A nucleic acid molecule comprising a T1R3 nucleotide sequence can alternatively be synthesized by use of the polymerase chain reaction (PCR) procedure, with the PCR oligonucleotide primers produced from the nucleotide sequences provided herein. See U.S. Patent Numbers 4,683,195 to Mullis *et al.* and 4,683,202 to Mullis. The PCR reaction provides a method for selectively increasing the concentration of a particular nucleic acid sequence even when that sequence has not been previously purified and is present only in a single copy in a

particular sample. The method can be used to amplify either single- or double-stranded DNA. The essence of the method involves the use of two oligonucleotide probes to serve as primers for the template-dependent, polymerase mediated replication of a desired nucleic acid molecule.

[0067] A wide variety of alternative cloning and *in vitro* amplification methodologies are well known to those skilled in the art. Examples of these techniques are found in, for example, Berger *et al.*, *Guide to Molecular Cloning Techniques*, METHODS IN ENZYMOLOGY 152, Academic Press, Inc., San Diego, CA (Berger), which is incorporated herein by reference in its entirety.

[0068] The polynucleotides of the invention may be used in hybridization techniques known to those skilled in the art, including but not limited to, Northern and Southern blotting and overgo hybridization (*see infra*). For example, polynucleotide probes of the invention may be used in tissue distribution studies and diagnostic assays. The T1R3 receptor of the invention is likely to be present and active in tissues other than those involved in taste perception. It is therefore likely that the feline T1R3 receptor serves multiple functions *in vivo*, such as, for example, regulation of amino acid metabolism in addition to taste perception.

[0069] Automated sequencing methods can be used to obtain or verify the nucleotide sequence of T1R3. The nucleotide sequences of the present invention are believed to be accurate. However, as is known in the art, nucleotide sequence obtained by automated methods may contain some errors. Nucleotide sequences determined by automation are typically at least about 90%, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of a given nucleic acid molecule. The actual sequence may be more precisely determined using manual sequencing methods, which are well known in the art. An error in a sequence which results in an insertion or deletion of one or more nucleotides may result in a frame shift in translation such that the predicted amino acid sequence will differ from that which would be predicted from the actual nucleotide sequence of the nucleic acid molecule, starting at the point of the mutation.

[0070] The nucleic acid molecules of the present invention, and fragments derived therefrom, are useful for screening for restriction fragment length polymorphism (RFLP) associated with certain disorders, as well as for genetic mapping.

[0071] The polynucleotide sequence information provided by the invention makes possible large-scale expression of the encoded polypeptide by techniques well known and routinely practiced in the art.

Vectors

[0072] Another aspect of the present invention is directed to vectors, or recombinant expression vectors, comprising any of the nucleic acid molecules described above. Vectors are used herein either to amplify DNA or RNA encoding T1R3 receptor and/or to express DNA which encodes T1R3 receptor. Examples of vectors include, but are not limited to, plasmids, phages, cosmids, episomes, viral particles or viruses, and integratable DNA fragments (*i.e.*, fragments integratable into the host genome by homologous recombination). Examples of viral particles include, but are not limited to, adenoviruses, baculoviruses, parvoviruses, herpesviruses, poxviruses, adeno-associated viruses, Semliki Forest viruses, vaccinia viruses, and retroviruses. Examples of expression vectors include, but are not limited to, pcDNA3 (Invitrogen) and pSVL (Pharmacia Biotech). Other expression vectors include, but are not limited to, pSPORT™ vectors, pGEM™ vectors (Promega), pPROEXvectors™ (LTI, Bethesda, MD), Bluescript™ vectors (Stratagene), pQE™ vectors (Qiagen), pSE420™ (Invitrogen), and pYES2™(Invitrogen).

[0073] Expression constructs may comprise T1R3-encoding polynucleotides operably linked to an endogenous or exogenous expression control DNA sequence and a transcription terminator. Expression control DNA sequences include promoters, enhancers, operators, and regulatory element binding sites generally, and are typically selected based on the expression systems in which the expression construct is to be utilized. Promoter and enhancer sequences are generally selected for the ability to increase gene expression, while operator sequences are generally selected for the ability to regulate gene expression. Expression constructs of the invention may also include sequences encoding one or more selectable markers that permit identification of host cells bearing the construct. Expression constructs may also include sequences that facilitate, or promote, homologous recombination in a host cell. Constructs of the invention also may include sequences necessary for replication in a host cell.

[0074] Expression constructs may be utilized for production of an encoded protein, but may also be utilized simply to amplify a T1R3-encoding polynucleotide sequence. In some embodiments, the vector is an expression vector wherein a polynucleotide of the invention is operably linked to a polynucleotide comprising an expression control sequence. Autonomously

replicating recombinant expression constructs such as plasmid and viral DNA vectors incorporating polynucleotides of the invention are also provided. Some expression vectors are replicable DNA constructs in which a DNA sequence encoding a T1R3 receptor is operably linked or connected to suitable control sequence(s) capable of effecting the expression of the receptor in a suitable host. Amplification vectors do not require expression control domains, but rather need only the ability to replicate in a host, such as conferred by an origin of replication, and a selection gene to facilitate recognition of transformants. The need for control sequences in the expression vector will vary depending upon the host selected and the transformation method chosen. Control sequences include a transcriptional promoter, an optional operator sequence to control transcription, a sequence encoding suitable mRNA ribosomal binding, and sequences which control the termination of transcription and translation.

[0075] Vectors of the invention may contain a promoter that is recognized by the host organism. The promoter sequences of the present invention may be prokaryotic, eukaryotic, or viral. Examples of suitable prokaryotic sequences include the P_R and P_L promoters of bacteriophage lambda (THE BACTERIOPHAGE LAMBDA, Hershey, A. D., Ed., Cold Spring Harbor Press, Cold Spring Harbor, NY (1973), which is incorporated herein by reference in its entirety; LAMBDA II, Hendrix, R. W., Ed., Cold Spring Harbor Press, Cold Spring Harbor, NY (1980), which is incorporated herein by reference in its entirety), the trp, recA, heat shock, and lacZ promoters of *E. coli,* and the SV40 early promoter (Benoist *et al. Nature*, 1981, *290*, 304-310), which is incorporated herein by reference in its entirety. Additional promoters include, but are not limited to, mouse mammary tumor virus, long terminal repeat of human immunodeficiency virus, maloney virus, cytomegalovirus immediate early promoter, Epstein Barr virus, Rous sarcoma virus, human actin, human myosin, human hemoglobin, human muscle creatine, and human metallothionein.

[0076] Additional regulatory sequences can also be included in vectors of the invention. Examples of suitable regulatory sequences are represented by the Shine-Dalgarno of the replicase gene of the phage MS-2 and of the gene cII of bacteriophage lambda. The Shine-Dalgarno sequence may be directly followed by DNA encoding a T1R3 receptor, resulting in the expression of the mature protein.

[0077] Moreover, suitable expression vectors can include an appropriate marker that allows the screening of transformed host cells. The transformation of the selected host is carried out using any one of the various techniques well known to the expert in the art and described in Sambrook *et al.*, *supra*.

[0078] An origin of replication or autonomously replicating sequence (ARS) can also be provided either by construction of the vector to include an exogenous origin or may be provided by the host cell chromosomal replication mechanism. If the vector is integrated into the host cell chromosome, the latter may be sufficient. Alternatively, rather than using vectors which contain viral origins of replication, one skilled in the art can transform mammalian cells by the method of co-transformation with a selectable marker and T1R3 DNA. An example of a suitable marker is dihydrofolate reductase (DHFR) or thymidine kinase (*see*, U.S. Patent No. 4,399,216).

[0079] Additional regulatory sequences that may be included in the polynucleotides of the invention include secretion signals which allow the encoded polypeptide to cross and/or lodge in cell membranes, or be secreted from the cell.

[0080] Nucleotide sequences encoding T1R3 may be recombined with vector DNA in accordance with conventional techniques, including blunt-ended or staggered-ended termini for ligation, restriction enzyme digestion to provide appropriate termini, filling in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and ligation with appropriate ligases. Techniques for such manipulation are disclosed by Sambrook *et al.*, *supra* and are well known in the art. Methods for construction of mammalian expression vectors are disclosed in, for example, Okayama *et al.*, *Mol. Cell. Biol.*, 1983, *3*, 280, Cosman *et al.*, *Mol. Immunol.*, 1986, *23*, 935, Cosman *et al.*, *Nature*, 1984, *312*, 768, EP-A-0367566, and WO 91/18982, each of which is incorporated herein by reference in its entirety.

Host cells

[0081] According to another aspect of the invention, host cells are provided, including prokaryotic and eukaryotic cells, comprising a polynucleotide of the invention (or vector of the invention) in a manner that permits expression of the encoded T1R3 polypeptide. Polynucleotides of the invention may be introduced into the host cell as part of a circular plasmid, or as linear DNA comprising an isolated protein-coding region or a viral vector. Methods for introducing DNA into the host cell that are well known and routinely practiced in the art include transformation, transfection, electroporation, nuclear injection, or fusion with carriers such as liposomes, micelles, ghost cells, and protoplasts. Expression systems of the invention include bacterial, yeast, fungal, plant, insect, invertebrate, vertebrate, and mammalian cell systems.

[0082] The invention provides host cells that are transformed or transfected (stably or transiently) with polynucleotides of the invention or vectors of the invention. As stated above, such host cells are useful for amplifying the polynucleotides and also for expressing a T1R3 polypeptide or fragment thereof encoded by the polynucleotide.

[0083] In still another related embodiment, the invention provides a method for producing a T1R3 polypeptide (or fragment thereof) comprising the steps of growing a host cell of the invention in a nutrient medium and isolating the polypeptide or variant thereof from the cell or the medium. Because the T1R3 receptor is a membrane-spanning polypeptide, it will be appreciated that, for some applications, such as certain activity assays, the preferable isolation may involve isolation of cell membranes containing the polypeptide embedded therein, whereas for other applications a more complete isolation may be preferable.

[0084] According to some aspects of the present invention, transformed host cells having an expression vector comprising any of the nucleic acid molecules described above are provided. Expression of the nucleotide sequence occurs when the expression vector is introduced into an appropriate host cell. Suitable host cells for expression of the polypeptides of the invention include, but are not limited to, prokaryotes, yeast, and eukaryotes. If a prokaryotic expression vector is employed, then the appropriate host cell would be any prokaryotic cell capable of expressing the cloned sequences. Suitable prokaryotic cells include, but are not limited to, bacteria of the genera *Escherichia*, *Bacillus*, *Salmonella*, *Pseudomonas*, *Streptomyces*, and *Staphylococcus*.

[0085] If a eukaryotic expression vector is employed, then the appropriate host cell would be any eukaryotic cell capable of expressing the cloned sequence. Eukaryotic cells may be cells of higher eukaryotes. Suitable eukaryotic cells include, but are not limited to, non-human mammalian tissue culture cells and human tissue culture cells. Host cells include, but are not limited to, insect cells, HeLa cells, Chinese hamster ovary cells (CHO cells), African green monkey kidney cells (COS cells), human HEK-293 cells, and murine 3T3 fibroblasts. Propagation of such cells in cell culture has become a routine procedure (*see*, TISSUE CULTURE, Academic Press, Kruse and Patterson, eds. (1973), which is incorporated herein by reference in its entirety).

[0086] In addition, a yeast host may be employed as a host cell. Yeast cells include, but are not limited to, the genera *Saccharomyces*, *Pichia*, and *Kluveromyces*. Yeast hosts may be *S. cerevisiae* and *P. pastoris*. Yeast vectors may contain an origin of replication sequence from a

2T yeast plasmid, an autonomously replication sequence (ARS), a promoter region, sequences for polyadenylation, sequences for transcription termination, and a selectable marker gene. Shuttle vectors for replication in both yeast and *E. coli* are also included herein.

[0087] Alternatively, insect cells may be used as host cells. In some embodiments, the polypeptides of the invention are expressed using a baculovirus expression system (*see*, Luckow *et al.*, *Bio/Technology*, 1988, *6*, 47; BACULOVIRUS EXPRESSION VECTORS: A LABORATORY MANUAL, O'Reilly *et al.* (Eds.), W.H. Freeman and Company, New York, 1992; and U.S. Patent No. 4,879,236, each of which is incorporated herein by reference in its entirety). In addition, the MAXBAC™ complete baculovirus expression system (Invitrogen) can, for example, be used for production in insect cells.

[0088] Host cells of the invention are a valuable source of immunogen for development of antibodies specifically immunoreactive with the T1R3 receptor. Host cells of the invention also are useful in methods for the large-scale production of T1R3 polypeptides wherein the cells are grown in a suitable culture medium and the desired polypeptide products are isolated from the cells, or from the medium in which the cells are grown, by purification methods known in the art, *e.g.*, conventional chromatographic methods including immunoaffinity chromatography, receptor affinity chromatography, hydrophobic interaction chromatography, lectin affinity chromatography, size exclusion filtration, cation or anion exchange chromatography, high pressure liquid chromatography (HPLC), reverse phase HPLC, and the like. Still other methods of purification include those methods wherein the desired protein is expressed and purified as a fusion protein having a specific tag, label, or chelating moiety that is recognized by a specific binding partner or agent. The purified protein can be cleaved to yield the desired protein, or can be left as an intact fusion protein. Cleavage of the fusion component may produce a form of the desired protein having additional amino acid residues as a result of the cleavage process.

[0089] Knowledge of the feline T1R3 nucleotide sequence allows for modification of cells to permit, or increase, expression of endogenous receptor. Cells can be modified (*e.g.*, by homologous recombination) to provide increased expression by replacing, in whole or in part, the naturally occurring T1R3 promoter with all or part of a heterologous promoter so that the cells express the receptor at higher or lower levels. The heterologous promoter is inserted in such a manner that it is operably linked to endogenous T1R3 coding sequence. (See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955.) It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (*e.g.*, ada, dhfr, and

the multifunctional CAD gene which encodes carbamoyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the T1R3 coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the T1R3 coding sequences in the cells.

Knock-out and transplacement animals

[0090] The DNA sequence information provided by the present invention also makes possible the development (*e.g.*, by homologous recombination strategies; see Capecchi, *Science 244*:1288-1292 (1989), which is incorporated herein by reference) of transgenic animals, including, for example, animals that fail to express functional T1R3 ("knock-out") or that express a variant thereof ("transplacement'). Such animals (especially small laboratory animals such as rats, rabbits, mice, and cats) are useful as models for studying the *in vivo* activities of T1R3 and modulators of T1R3.

Antisense and siRNA

[0091] Also encompassed by the invention are antisense and short interfering polynucleotides that recognize and hybridize to polynucleotides encoding T1R3. Full-length and fragment antisense polynucleotides are provided. Fragment antisense molecules of the invention include (i) those that specifically recognize and hybridize to T1R3 RNA (as determined by sequence comparison of DNA encoding T1R3 to DNA encoding other known molecules). Identification of sequences unique to T1R3 encoding polynucleotides can be deduced through use of any publicly available sequence database, and/or through use of commercially available sequence comparison programs. After identification of the desired sequences, isolation through restriction digestion or amplification using any of the various polymerase chain reaction techniques well known in the art can be performed. Antisense polynucleotides are particularly relevant to regulation of expression of T1R3 receptor by those cells expressing T1R3 mRNA.

[0092] Antisense nucleic acids (preferably 10 to 30 base-pair oligonucleotides) capable of specifically binding to T1R3 expression control sequences or T1R3 RNA are introduced into cells (*e.g.*, by a viral vector or colloidal dispersion system such as a liposome). The antisense nucleic acid binds to the target nucleotide sequence in the cell and prevents transcription and/or translation of the target sequence. Phosphorothioate and methylphosphonate antisense oligonucleotides are specifically contemplated for therapeutic use by the invention. Locked

nucleic acids are also specifically contemplated for therapeutic use by the present invention. (See, for example, Wahlestedt *et al.*, *Proc. Natl. Acad. Sci. USA*, 97(10), 5633-5638 (2000), which is incorporated by reference in its entirety) The antisense oligonucleotides may be further modified by adding poly-L-lysine, transferrin polylysine, or cholesterol moieties at their 5' end. Suppression of T1R3 expression at either the transcriptional or translational level is useful to generate cellular or animal models for diseases/conditions characterized by aberrant T1R3 expression.

[0093] Antisense oligonucleotides, or fragments of nucleotide sequence of SEQ ID NO:1, or sequences complementary or homologous thereto, derived from the nucleotide sequences of the present invention encoding T1R3 receptor are useful as diagnostic tools for probing gene expression in various tissues. For example, tissue can be probed *in situ* with oligonucleotide probes carrying detectable groups by conventional autoradiography techniques to investigate native expression of this enzyme or pathological conditions relating thereto. Antisense oligonucleotides may be directed to regulatory regions of a T1R3 nucleotide sequence, or mRNA corresponding thereto, including, but not limited to, the initiation codon, TATA box, enhancer sequences, and the like.

[0094] Those of skill in the art recognize that the antisense oligonucleotides that inhibit the expression and/or biological activity of a T1R3 receptor may be predicted using any gene encoding a T1R3 receptor. Specifically, antisense nucleic acid molecules comprise a sequence preferably complementary to at least about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, 250 or 500 nucleotides or an entire T1R3 receptor gene sequence. The antisense oligonucleotides may comprise a sequence complementary to about 15 consecutive nucleotides of the coding strand of the T1R3 receptor-encoding sequence.

[0095] In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding a T1R3 protein. The coding strand may also include regulatory regions of the T1R3 sequence. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding a T1R3 protein. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions (UTR)).

[0096] Antisense oligonucleotides may be directed to regulatory regions of a nucleotide sequence encoding a T1R3 protein, or mRNA corresponding thereto, including, but not limited to, the initiation codon, TATA box, enhancer sequences, and the like. Given the coding strand sequences provided herein, antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of a T1R3 mRNA, but also may be an oligonucleotide that is antisense to only a portion of the coding or noncoding region of the mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of an mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length.

[0097] Another means to inhibit the activity of a T1R3 receptor according to the invention is via RNA interference (RNAi) (*see e.g.*, Elbashir *et al.*, *Nature,* 411:494-498 (2001); Elbashir *et al.*, *Genes Development*, 15:188-200 (2001)). RNAi is the process of sequence-specific, post-transcriptional gene silencing, initiated by double-stranded RNA (dsRNA) that is homologous in sequence to the silenced gene (*e.g.*, is homologous in sequence to the sequence of a T1R3, for example but not limited to the sequence as set forth in SEQ ID NO:1). siRNA-mediated silencing is thought to occur post-transcriptionally and/or transcriptionally. For example, siRNA duplexes may mediate post-transcriptional gene silencing by reconstitution of siRNA-protein complexes (siRNPs), which guide mRNA recognition and targeted cleavage.

[0098] Accordingly, another form of a T1R3 inhibitory compound of the invention is a short interfering RNA (siRNA) directed against a T1R3-encoding sequence. Exemplary siRNAs are siRNA duplexes (for example, 10-25, preferably 20, 21, 22, 23, 24, or 25 residues in length) having a sequence homologous or identical to a fragment of the T1R3 sequence set forth as SEQ ID NO:1 and having a symmetric 2-nucleotide 3'-overhang. The 2-nucleotide 3' overhang may be composed of (2'-deoxy) thymidine because it reduces costs of RNA synthesis and may enhance nuclease resistance of siRNAs in the cell culture medium and within transfected cells. Substitution of uridine by thymidine in the 3' overhang is also well tolerated in mammalian cells, and the sequence of the overhang appears not to contribute to target recognition.

Polypeptides

[0099] The invention also provides purified and isolated mammalian T1R3 receptor polypeptides encoded by a polynucleotide of the invention. Some embodiments include a feline T1R3 polypeptide comprising the amino acid sequence of SEQ ID NO:2, or fragments thereof

comprising an epitope specific to the polypeptide. A reference to "epitope specific to" or "polypeptide-specific epitope," or variations thereof, indicates that a portion of the T1R3 receptor or amino acid sequence is recognizable by an antibody that is specific for the T1R3 or amino acid sequence.

[0100] Included within the scope of the invention are polypeptides encoded by feline allelic variants of T1R3. The allelic variants of the T1R3 receptor of the invention may modify the taste perception of a mammal, such as a cat, to a taste stimulus. Such functional amino acid sequence modifications may account for differences in intraspecies (*e.g.*, breed-specific) taste perception.

[0101] Extracellular epitopes are useful for generating and screening for antibodies and other binding compounds that bind to a T1R3 receptor. Thus, in another embodiment, the invention provides a purified and isolated polypeptide comprising at least one extracellular domain of the T1R3 receptor. Also included is a polypeptide comprising a T1R3 receptor fragment selected from the group consisting of an extracellular domain of T1R3 (residues 1-571, 628-641, 705-730, and 787-794 of SEQ ID NO:2), a transmembrane domain of T1R3 (residues 572-594, 610-627, 642-664, 681-704, 731-754, 767-780, and 795-812 of SEQ ID NO:2), and an intracellular domain of T1R3 (residues 595-609, 665-680, 755-766, and 813-865 of SEQ ID NO:2). Polypeptide fragments of the invention may be continuous portions of the native receptor. However, it will also be appreciated that knowledge of the T1R3 gene and protein sequences as provided herein permits recombination of various domains that are not contiguous in the native protein.

[0102] The invention embraces polypeptides that preferably have at least 99%, at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 74%, at least 73%, at least 72%, at least 71%, at least 70%, at least 65%, at least 60%, at least 55% or at least 50% identity and/or homology to the polypeptide of the invention.

[0103] Polypeptides of the invention may be isolated from natural cell sources or may be chemically synthesized, but are preferably produced by recombinant procedures involving host cells of the invention. Use of mammalian host cells is expected to provide for such post-translational modifications (*e.g.*, glycosylation, truncation, lipidation, and phosphorylation) as may be needed to confer optimal biological activity on recombinant expression products of the invention.

[0104] The invention also embraces variant T1R3 polypeptides. In one example, insertion variants are provided wherein one or more amino acid residues supplement a T1R3 amino acid sequence such as SEQ ID NO:2. Insertions may be located at either or both termini of the protein, or may be positioned within internal regions of the amino acid sequence. Insertional variants with additional residues at either or both termini can include, for example, fusion proteins and proteins including amino acid tags or labels.

[0105] Insertion variants include T1R3 polypeptides wherein one or more amino acid residues are added to a biologically active fragment thereof.

[0106] The invention also embraces T1R3 variants having additional amino acid residues that result from use of specific expression systems. For example, use of commercially available vectors that express a desired polypeptide as part of a glutathione-S-transferase (GST) fusion product provides the desired polypeptide having an additional glycine residue at position -1 after cleavage of the GST component from the desired polypeptide. Variants that result from expression in other vector systems are also contemplated.

[0107] In another aspect, the invention provides deletion variants wherein one or more amino acid residues in a T1R3 polypeptide are removed. Deletions can be effected at one or both termini of the T1R3 polypeptide, or with removal of one or more non-terminal amino acid residues of T1R3. Deletion variants, therefore, include all fragments of a T1R3 polypeptide.

[0108] The invention also embraces polypeptide fragments that maintain biological (*e.g.*, ligand binding) and/or immunological properties of a T1R3 polypeptide.

[0109] As used in the present invention, polypeptide fragments preferably comprise at least 10, 15, 20, 25, 30, 35, 40, 45, or 50 consecutive amino acids of SEQ ID NO:2. Some polypeptide fragments display antigenic properties unique to, or specific for, feline T1R3. Fragments of the invention having the desired biological and immunological properties can be prepared by any of the methods well known and routinely practiced in the art.

[0110] In still another aspect, the invention provides substitution variants of T1R3 polypeptides. Substitution variants include those polypeptides wherein one or more amino acid residues of a T1R3 polypeptide are removed and replaced with alternative residues. In one aspect, the substitutions are conservative in nature; however, the invention embraces substitutions that are also non-conservative. Conservative substitutions for this purpose may be defined as set out in Tables 1, 2, or 3 below.

[0111] Variant polypeptides include those wherein conservative substitutions have been introduced by modification of polynucleotides encoding polypeptides of the invention. Amino acids can be classified according to physical properties and contribution to secondary and tertiary protein structure. A conservative substitution is recognized in the art as a substitution of one amino acid for another amino acid that has similar properties. Exemplary conservative substitutions are set out in Table 1 (from WO 97/09433, page 10, published March 13, 1997 (PCT/GB96/02197, filed 9/6/96), immediately below.

Table 1

Conservative Substitutions I

SIDE CHAIN CHARACTERISTIC	AMINO ACID
Aliphatic	
Non-polar	G A P
	I L V
Polar - uncharged	C S T M
	N Q
Polar - charged	D E
	K R
Aromatic	H F W Y
Other	N Q D E

Alternatively, conservative amino acids can be grouped as described in Lehninger, [BIOCHEMISTRY, Second Edition; Worth Publishers, Inc. NY, NY (1975), pp.71-77] as set out in Table 2, below.

Table 2
Conservative Substitutions II

SIDE CHAIN CHARACTERISTIC	AMINO ACID
Non-polar (hydrophobic)	
A. Aliphatic:	A L I V P
B. Aromatic:	F W
C. Sulfur-containing:	M
D. Borderline:	G
Uncharged-polar	
A. Hydroxyl:	S T Y
B. Amides:	N Q
C. Sulfhydryl:	C
D. Borderline:	G
Positively Charged (Basic):	K R H
Negatively Charged (Acidic):	D E

As still another alternative, exemplary conservative substitutions are set out in Table 3, below.

Table 3

Conservative Substitutions III

Original Residue	Exemplary Substitution
Ala (A)	Val, Leu, Ile
Arg (R)	Lys, Gln, Asn
Asn (N)	Gln, His, Lys, Arg
Asp (D)	Glu
Cys (C)	Ser
Gln (Q)	Asn
Glu (E)	Asp
His (H)	Asn, Gln, Lys, Arg
Ile (I)	Leu, Val, Met, Ala, Phe,
Leu (L)	Ile, Val, Met, Ala, Phe
Lys (K)	Arg, Gln, Asn
Met (M)	Leu, Phe, Ile
Phe (F)	Leu, Val, Ile, Ala
Pro (P)	Gly
Ser (S)	Thr
Thr (T)	Ser
Trp (W)	Tyr
Tyr (Y)	Trp, Phe, Thr, Ser
Val (V)	Ile, Leu, Met, Phe, Ala

[0112] It should be understood that the definition of polypeptides of the invention is intended to include polypeptides bearing modifications other than insertion, deletion, or substitution of amino acid residues. By way of example, the modifications may be covalent in nature, and include for example, chemical bonding with polymers, lipids, other organic, and inorganic moieties. Such derivatives may be prepared to increase circulating half-life of a polypeptide, or may be designed to improve the targeting capacity of the polypeptide for desired cells, tissues, or organs. Similarly, the invention further embraces T1R3 polypeptides that have been covalently modified to include one or more water-soluble polymer attachments such as polyethylene glycol, polyoxyethylene glycol, or polypropylene glycol. Variants that display ligand binding properties of native T1R3 and are expressed at higher levels, as well as variants that provide for constitutively active receptors, are particularly useful in assays of the invention; the variants are also useful in providing cellular, tissue and animal models of diseases/conditions characterized by aberrant T1R3 activity.

[0113] In a related embodiment, the present invention provides compositions comprising purified polypeptides of the invention. Some compositions comprise, in addition to the

polypeptide of the invention, a pharmaceutically acceptable (*i.e.*, sterile and non-toxic) liquid, semisolid, or solid diluent that serves as a pharmaceutical vehicle, excipient, or medium. Any diluent known in the art may be used. Exemplary diluents include, but are not limited to, water, saline solutions, polyoxyethylene sorbitan monolaurate, magnesium stearate, methyl- and propylhydroxybenzoate, talc, alginates, starches, lactose, sucrose, dextrose, sorbitol, mannitol, glycerol, calcium phosphate, mineral oil, and cocoa butter.

[0114] Variants that display ligand-binding properties of native T1R3 and are expressed at higher levels, as well as variants that provide for constitutively active receptors, are particularly useful in assays of the invention; the variants are also useful in assays of the invention and in providing cellular, tissue and animal models of diseases/conditions characterized by aberrant T1R3 activity.

Antibodies

[0115] Also included in the present invention are antibodies (*e.g.*, monoclonal and polyclonal antibodies, single chain antibodies, chimeric antibodies, bifunctional/bispecific antibodies, humanized antibodies, human antibodies, felinized antibodies, feline antibodies, and complementary determining region (CDR)-grafted antibodies, including compounds which include CDR sequences which specifically recognize a polypeptide of the invention) specific for T1R3 receptor or fragments thereof. Antibody fragments, including Fab, Fab′, $F(ab')_2$, and F_v, are also provided by the invention. The term "specific for," when used to describe antibodies of the invention, indicates that the variable regions of the antibodies of the invention recognize and bind T1R3 polypeptides, preferably exclusively (*i.e.*, are able to distinguish T1R3 polypeptides from other known polypeptides by virtue of measurable differences in binding affinity, despite the possible existence of localized sequence identity, homology, or similarity between T1R3 and such polypeptides). It will be understood that specific antibodies may also interact with other proteins (for example, *S. aureus* protein A or other antibodies in ELISA techniques) through interactions with sequences outside the variable region of the antibodies, and, in particular, in the constant region of the molecule. Screening assays to determine binding specificity of an antibody of the invention are well known and routinely practiced in the art. For a comprehensive discussion of such assays, see Harlow *et al.* (Eds.), ANTIBODIES A LABORATORY MANUAL; Cold Spring Harbor Laboratory; Cold Spring Harbor, NY (1988), Chapter 6. Antibodies that recognize and bind fragments of the T1R3 polypeptides of the invention are also contemplated, provided that the antibodies are specific for T1R3 polypeptides. Antibodies of the invention can be produced using any method well known and routinely practiced in the art.

[0116] The invention provides an antibody that is specific for the feline T1R3 of the invention. Antibodies that can be generated from polypeptides that have previously been described in the literature and that are capable of fortuitously cross-reacting with T1R3 (*e.g.*, due to the fortuitous existence of a similar epitope in both polypeptides) are considered "cross-reactive" antibodies. Such cross-reactive antibodies are not antibodies that are "specific" for T1R3 receptor. The determination of whether an antibody is specific for T1R3 or is cross-reactive with another known receptor is made using any of several assays, such as Western blotting assays, that are well known in the art. For identifying cells that express T1R3 and also for modulating T1R3-ligand binding activity, antibodies that specifically bind to an extracellular epitope of the T1R3 may be used.

[0117] In some variations, the invention provides monoclonal antibodies. Hybridomas that produce such antibodies also are intended as aspects of the invention. In yet another variation, the invention provides a felinized antibody. Felinized antibodies are useful for *in vivo* therapeutic indications.

[0118] In another variation, the invention provides a cell-free composition comprising polyclonal antibodies, wherein at least one of the antibodies is an antibody of the invention specific for T1R3. Antisera isolated from an animal is an exemplary composition, as is a composition comprising an antibody fraction of an antisera that has been resuspended in water or in another diluent, excipient, or carrier.

[0119] In still another related embodiment, the invention provides an anti-idiotypic antibody specific for an antibody that is specific for T1R3.

[0120] It is well known that antibodies contain relatively small antigen binding domains that can be isolated chemically or by recombinant techniques. Such domains are useful T1R3 binding molecules themselves, and also may be reintroduced into other antibodies or fused to toxins or other polypeptides. Thus, in still another embodiment, the invention provides a polypeptide comprising a fragment of a T1R3-specific antibody, wherein the fragment and the polypeptide bind to the T1R3 receptor. By way of non-limiting example, the invention provides polypeptides that are single chain antibodies and CDR-grafted antibodies.

[0121] Non-feline antibodies may be felinized by any of the methods known in the art. In one method, the non-feline CDRs are inserted into a feline antibody or consensus antibody framework sequence. Similarly, non-human antibodies may be humanized by methods known in

the art. In one embodiment, non-human CDRs are inserted into a human antibody or consensus antibody framework sequence. Further changes can then be introduced into the antibody framework to modulate affinity or immunogenicity.

[0122] Antibodies of the invention are useful for, *e.g.*, therapeutic purposes (such as by modulating activity of T1R3), diagnostic purposes (such as detecting or quantitating T1R3 activity), and also for purification of T1R3. Kits comprising an antibody of the invention for any of the purposes described herein are also included within the scope of the invention. In general, a kit of the invention preferably includes a control antigen for which the antibody is immunospecific.

Compositions

[0123] Mutations in the T1R3 gene that result in loss of normal function of the T1R3 gene product underlie some T1R3-related disease states. The invention comprehends gene and peptide therapy, for example, to restore T1R3 activity to treat those disease states. Delivery of a functional T1R3 gene to appropriate cells is effected *ex vivo*, *in situ*, or *in vivo* by use of vectors, and more particularly viral vectors (*e.g.*, adenovirus, adeno-associated virus, or a retrovirus), or *ex vivo* by use of physical DNA transfer methods (*e.g.*, liposomes or chemical treatments). See, for example, Anderson, *Nature*, supplement to vol. 392, No. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, *Science, 244*: 1275-1281 (1989); Verma, *Scientific American*: 68-84 (1990); and Miller, *Nature, 357*: 455-460 (1992). Alternatively, it is contemplated that in other disease states, preventing the expression of, or inhibiting the activity of, T1R3 will be useful in treatment. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of T1R3.

[0124] Another aspect of the present invention is directed to compositions, including pharmaceutical compositions, comprising any of the nucleic acid molecules or recombinant expression vectors described above and an acceptable carrier or diluent. The carrier or diluent may be pharmaceutically acceptable. Suitable carriers are described in the most recent edition of *Remington's Pharmaceutical Sciences*, A. Osol, a standard reference text in this field, which is incorporated herein by reference in its entirety. Examples of such carriers or diluents include, but are not limited to, water, saline, Ringer's solution, dextrose solution, and 5% serum albumin. Liposomes and nonaqueous vehicles such as fixed oils may also be used. The formulations may be sterilized by commonly used techniques.

[0125] Also within the scope of the invention are compositions comprising polypeptides, polynucleotides, or antibodies of the invention that have been formulated with, *e.g.*, a pharmaceutically acceptable carrier.

[0126] The invention also provides methods of using antibodies of the invention. For example, the invention provides a method for modulating ligand-binding of a T1R3 receptor comprising the step of contacting the receptor with an antibody specific for the T1R3 polypeptide, under conditions wherein the antibody binds the receptor.

Methods of identifying ligands and modulators

[0127] The invention also provides assays to identify compounds that bind and/or modulate T1R3. A "T1R3 binding partner" is a compound that directly or indirectly binds a T1R3 polypeptide of the invention. One assay of the invention comprises the steps of: (a) contacting T1R3 with a compound suspected of binding T1R3; and (b) measuring binding between the compound and T1R3. In one variation, the composition comprises a cell expressing T1R3 on its surface. In another variation, isolated T1R3 or cell membranes comprising T1R3 are employed. The binding may be measured directly, *e.g.*, by using a labeled compound, or may be measured indirectly. Compounds identified as binding T1R3 receptor may be further tested in other assays including, but not limited to, T1R3 activity assays and/or *in vivo* models, in order to confirm or quantitate their activity.

[0128] Specific binding molecules, including natural ligands and synthetic compounds, can be identified or developed using isolated or recombinant T1R3 products, T1R3 variants, or preferably, cells expressing such products. Binding partners are useful for purifying T1R3 products and detection or quantification of T1R3 products in fluid and tissue samples using known immunological procedures. Binding molecules are also manifestly useful in modulating (*i.e.*, blocking, inhibiting or stimulating) biological activities of T1R3, especially those activities involved in signal transduction.

[0129] The DNA and amino acid sequence information provided by the present invention also makes possible identification of binding partner compounds with which a T1R3 polypeptide or polynucleotide will interact. Methods to identify binding partner compounds include solution assays, *in vitro* assays wherein T1R3 polypeptides are immobilized, and cell-based assays. Identification of binding partner compounds of T1R3 polypeptides provides candidates for

therapeutic or prophylactic intervention in pathologies associated with T1R3 normal and aberrant biological activity.

[0130] The invention includes several assay systems for identifying T1R3-binding partners. In solution assays, methods of the invention comprise the steps of (a) contacting a T1R3 polypeptide with one or more candidate binding partner compounds and (b) identifying the compounds that bind to the T1R3 polypeptide. Identification of the compounds that bind the T1R3 polypeptide can be achieved by isolating the T1R3 polypeptide/binding partner complex, and separating the binding partner compound from the T1R3 polypeptide. An additional step of characterizing the physical, biological, and/or biochemical properties of the binding partner compound is also comprehended in another embodiment of the invention. In one aspect, the T1R3 polypeptide/binding partner complex is isolated using an antibody immunospecific for either the T1R3 polypeptide or the candidate binding partner compound.

[0131] In still other embodiments, either the T1R3 polypeptide or the candidate binding partner compound comprises a label or tag that facilitates its isolation, and methods of the invention to identify binding partner compounds include a step of isolating the T1R3 polypeptide/binding partner complex through interaction with the label or tag. An exemplary tag of this type is a poly-histidine sequence, generally around six histidine residues, that permits isolation of a compound so labeled using nickel chelation. Other labels and tags, such as the FLAG® tag (Eastman Kodak, Rochester, NY), well known and routinely used in the art, are embraced by the invention.

[0132] In one variation of an *in vitro* assay, the invention provides a method comprising the steps of (a) contacting an immobilized T1R3 polypeptide with a candidate binding partner compound and (b) detecting binding of the candidate compound to the T1R3 polypeptide. In an alternative embodiment, the candidate binding partner compound is immobilized and binding of T1R3 is detected. Immobilization is accomplished using any of the methods well known in the art, including covalent bonding to a support, a bead, or a chromatographic resin, as well as non-covalent, high affinity interactions such as antibody binding, or use of streptavidin/biotin binding wherein the immobilized compound includes a biotin moiety. The support may, for example, be formulated into a feline-specific electronic tongue. Detection of binding can be accomplished (i) using a radioactive label on the compound that is not immobilized, (ii) using a fluorescent label on the non-immobilized compound, (iii) using an antibody immunospecific for the non-immobilized compound, (iv) using a label on the non-immobilized compound that excites a

fluorescent support to which the immobilized compound is attached, as well as other techniques well known and routinely practiced in the art.

[0133] The invention also provides cell-based assays to identify binding partner compounds of a T1R3 polypeptide. In one embodiment, the invention provides a method comprising the steps of contacting a T1R3 polypeptide expressed on the surface of a cell with a candidate binding partner compound and detecting binding of the candidate binding partner compound to the T1R3 polypeptide. In some embodiments, the detection comprises detecting physiological event in the cell caused by the binding of the molecule.

[0134] Another aspect of the present invention is directed to methods of identifying compounds that bind to either T1R3 or nucleic acid molecules encoding T1R3, comprising contacting T1R3, or a nucleic acid molecule encoding the same, with a compound, and determining whether the compound binds T1R3 or a nucleic acid molecule encoding the same. Binding can be determined by binding assays which are well known to the skilled artisan, including, but not limited to, gel-shift assays, Western blots, radiolabeled competition assay, phage-based expression cloning, co-fractionation by chromatography, co-precipitation, cross-linking, interaction trap/two-hybrid analysis, southwestern analysis, ELISA, and the like, which are described in, for example, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, 1999, John Wiley & Sons, NY, which is incorporated herein by reference in its entirety. The compounds to be screened include (which may include compounds which are suspected to bind T1R3, or a nucleic acid molecule encoding the same), but are not limited to, extracellular, intracellular, biological, or chemical origin. The methods of the invention also embrace ligands, especially neuropeptides, that are attached to a label, such as a radiolabel (*e.g.*, ^{125}I, ^{35}S, ^{32}P, ^{33}P, ^{3}H), a fluorescence label, a chemiluminescent label, an enzymic label, and an immunogenic label. Modulators falling within the scope of the invention include, but are not limited to, non-peptide molecules such as non-peptide mimetics, non-peptide allosteric effectors, and peptides. The T1R3 polypeptide or polynucleotide employed in such a test may either be free in solution, attached to a solid support, borne on a cell surface or located intracellularly, or associated with a portion of a cell. One skilled in the art can, for example, measure the formation of complexes between T1R3 and the compound being tested. Alternatively, one skilled in the art can examine the diminution in complex formation between T1R3 and its substrate caused by the compound being tested. In some embodiments of the invention, the recognition sites of the T1R3 receptor are coupled with a monitoring system, either electrical or optical. An appropriate chemical stimulus can bind to the receptor's ligand binding domain, changing the receptor conformation to

a degree that the coupled electronics or optical changes can be observed on a read-out. Such a device could be developed into a feline-specific electronic tongue, for example.

[0135] In another embodiment of the invention, high throughput screening for compounds having suitable binding affinity to T1R3 receptor is employed. Briefly, large numbers of different small peptide test compounds are synthesized on a solid substrate. The peptide test compounds are contacted with T1R3 and washed. Bound T1R3 is then detected by methods well known in the art. Purified polypeptides of the invention can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture the protein and immobilize it on the solid support.

[0136] Generally, an expressed T1R3 receptor can be used for HTS binding assays in conjunction with a ligand, such as an amino acid or carbohydrate. The identified peptide is labeled with a suitable radioisotope, including, but not limited to, ^{125}I, ^{3}H, ^{35}S or ^{32}P, by methods that are well known to those skilled in the art. Alternatively, the peptides may be labeled by well-known methods with a suitable fluorescent derivative (Baindur *et al.*, *Drug Dev. Res.*, 1994, *33*, 373-398; Rogers, *Drug Discovery Today*, 1997, *2*, 156-160). Radioactive ligand specifically bound to the receptor in membrane preparations made from the cell line expressing the recombinant protein can be detected in HTS assays in one of several standard ways, including filtration of the receptor-ligand complex to separate bound ligand from unbound ligand (Williams, *Med. Res. Rev.*, 1991, *11*, 147-184; Sweetnam *et al.*, *J. Natural Products*, 1993, *56*, 441-455). Alternative methods include a scintillation proximity assay (SPA) or a FlashPlate format in which such separation is unnecessary (Nakayama, *Cur. Opinion Drug Disc. Dev.*, 1998, *1*, 85-91; Bossé *et al.*, *J. Biomolecular Screening*, 1998, *3*, 285-292.). Binding of fluorescent ligands can be detected in various ways, including fluorescence energy transfer (FRET), direct spectrophotofluorometric analysis of bound ligand, or fluorescence polarization (Rogers, *Drug Discovery Today*, 1997, *2*, 156-160; Hill, *Cur. Opinion* Drug *Disc. Dev.*, 1998, *1*, 92-97).

[0137] Other assays may be used to identify specific ligands of a T1R3 receptor, including assays that identify ligands of the target protein through measuring direct binding of test ligands to the target protein, as well as assays that identify ligands of target proteins through affinity ultrafiltration with ion spray mass spectroscopy/HPLC methods or other physical and analytical methods. Alternatively, such binding interactions are evaluated indirectly using the yeast two-hybrid system described in Fields *et al.*, *Nature,* 340:245-246 (1989), and Fields *et al.*, *Trends in Genetics*, 10:286-292 (1994), both of which are incorporated herein by reference. The two-

hybrid system is a genetic assay for detecting interactions between two proteins or polypeptides. It can be used to identify proteins that bind to a known protein of interest, or to delineate domains or residues critical for an interaction. Variations on this methodology have been developed to clone genes that encode DNA binding proteins, to identify peptides that bind to a protein, and to screen for drugs. The two-hybrid system exploits the ability of a pair of interacting proteins to bring a transcription activation domain into close proximity with a DNA binding domain that binds to an upstream activation sequence (UAS) of a reporter gene, and is generally performed in yeast. The assay requires the construction of two hybrid genes encoding (1) a DNA-binding domain that is fused to a first protein and (2) an activation domain fused to a second protein. The DNA-binding domain targets the first hybrid protein to the UAS of the reporter gene; however, because most proteins lack an activation domain, this DNA-binding hybrid protein does not activate transcription of the reporter gene. The second hybrid protein, which contains the activation domain, cannot by itself activate expression of the reporter gene because it does not bind the UAS. However, when both hybrid proteins are present, the noncovalent interaction of the first and second proteins tethers the activation domain to the UAS, activating transcription of the reporter gene. For example, when the first protein is a receptor, or fragment thereof, that is known to interact with another protein or nucleic acid, this assay can be used to detect agents that interfere with the binding interaction. Expression of the reporter gene is monitored as different test agents are added to the system. The presence of an inhibitory agent results in lack of a reporter signal.

[0138] The yeast two-hybrid assay can also be used to identify proteins that bind to the gene product. In an assay to identify proteins that bind to a T1R3 receptor, or fragment thereof, a fusion polynucleotide encoding both a T1R3 receptor (or fragment) and a UAS binding domain (*i.e.*, a first protein) may be used. In addition, a large number of hybrid genes each encoding a different second protein fused to an activation domain are produced and screened in the assay. Typically, the second protein is encoded by one or more members of a total cDNA or genomic DNA fusion library, with each second protein-coding region being fused to the activation domain. This system is applicable to a wide variety of proteins, and it is not necessary to know the identity or function of the second binding protein. The system is highly sensitive and can detect interactions not revealed by other methods; even transient interactions may trigger transcription to produce a stable mRNA that can be repeatedly translated to yield the reporter protein.

[0139] Other assays may be used to search for agents that bind to the target protein. One such screening method to identify direct binding of test ligands to a target protein is described in U.S. Patent No. 5,585,277, incorporated herein by reference. This method relies on the principle that proteins generally exist as a mixture of folded and unfolded states, and continually alternate between the two states. When a test ligand binds to the folded form of a target protein (*i.e.*, when the test ligand is a ligand of the target protein), the target protein molecule bound by the ligand remains in its folded state. Thus, the folded target protein is present to a greater extent in the presence of a test ligand which binds the target protein, than in the absence of a ligand. Binding of the ligand to the target protein can be determined by any method that distinguishes between the folded and unfolded states of the target protein. The function of the target protein need not be known in order for this assay to be performed. Virtually any agent can be assessed by this method as a test ligand, including, but not limited to, metals, polypeptides, proteins, lipids, polysaccharides, polynucleotides and small organic molecules.

[0140] Another method for identifying ligands of a target protein is described in Wieboldt *et al.*, *Anal. Chem.*, 69:1683-1691 (1997), incorporated herein by reference. This technique screens combinatorial libraries of 20-30 agents at a time in solution phase for binding to the target protein. Agents that bind to the target protein are separated from other library components by simple membrane washing. The specifically selected molecules that are retained on the filter are subsequently liberated from the target protein and analyzed by HPLC and pneumatically assisted electrospray (ion spray) ionization mass spectroscopy. This procedure selects library components with the greatest affinity for the target protein, and is particularly useful for small molecule libraries.

[0141] Other embodiments of the invention comprise using competitive screening assays in which neutralizing antibodies capable of binding a polypeptide of the invention specifically compete with a test compound for binding to the polypeptide. In this manner, the antibodies can be used to detect the presence of any peptide that shares one or more antigenic determinants with T1R3 receptor. Radiolabeled competitive binding studies are described in A.H. Lin *et al.*, *Antimicrobial Agents and Chemotherapy*, 1997, 41(10): 2127-2131, the disclosure of which is incorporated herein by reference in its entirety.

[0142] Another aspect of the present invention is directed to methods of identifying compounds that modulate (*i.e.*, increase or decrease) activity of T1R3 comprising contacting T1R3 with a compound, and determining whether the compound modifies activity of T1R3. The activity in the presence of the test compound is compared to the activity in the absence of the test

compound. Where the activity of the sample containing the test compound is higher than the activity in the sample lacking the test compound, the compound is an agonist. Similarly, where the activity of the sample containing the test compound is lower than the activity in the sample lacking the test compound, the compound is an antagonist.

[0143] Agents that modulate (*i.e.*, increase, decrease, or block) T1R3 activity or expression also may be identified, for example, by incubating a putative modulator with a cell containing a T1R3 polypeptide or polynucleotide and determining the effect of the putative modulator on T1R3 activity or expression. The selectivity of a compound that modulates the activity of T1R3 can be evaluated by comparing its effects on T1R3 to its effect on other T1R receptors. Selective modulators may include, for example, antibodies and other proteins, peptides, or organic molecules that specifically bind to a T1R3 polypeptide or a T1R3-encoding nucleic acid. Modulators of T1R3 activity will be therapeutically useful in treatment of diseases and physiological conditions in which normal or aberrant T1R3 activity is involved. Compounds identified as modulating T1R3 activity may be further tested in other assays including, but not limited to, *in vivo* models, in order to confirm or quantitate their activity.

[0144] The invention also provides methods for identifying a T1R3 modulator by: (a) contacting a T1R3 binding partner and a composition comprising a T1R3 receptor in the presence and in the absence of a putative modulator compound; (b) detecting binding between the binding partner and the T1R3 receptor; and (c) identifying a putative modulator compound or a modulator compound in view of decreased or increased binding between the binding partner and the T1R3 receptor in the presence of the putative modulator, as compared to binding in the absence of the putative modulator. Compounds identified as modulators of binding between T1R3 and a T1R3 binding partner may be further tested in other assays including, but not limited to, *in vivo* models, in order to confirm or quantitate their activity.

[0145] The invention also includes within its scope high-throughput screening (HTS) assays to identify compounds that interact with, enhance, or inhibit biological activity (*i.e.*, affect enzymatic activity, binding activity, *etc.*) of a T1R3 polypeptide. HTS assays permit screening of large numbers of compounds in an efficient manner. Cell-based HTS systems are contemplated to investigate T1R3 receptor-ligand interaction. HTS assays are designed to identify "hits" or "lead compounds" having the desired property, from which modifications can be designed to improve the desired property. Chemical modification of the "hit" or "lead compound" is often based on an identifiable structure/activity relationship between the "hit" and the T1R3 polypeptide.

[0146] For example, modulators of T1R3 activity may be identified by expressing the T1R3 receptor in a heterologous cultured mammalian cell line, such as HEK cells, and detecting receptor activity in the presence and absence of a test compound by monitoring changes in intracellular calcium using a calcium-specific intracellular dye. In another embodiment, this process may be automated using a high-throughput screening device.

[0147] Candidate modulators contemplated by the invention include compounds selected from libraries of either potential activators or potential inhibitors. There are a number of different libraries used for the identification of small molecule modulators, including: (1) chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides, or organic molecules. Chemical libraries consist of random chemical structures, some of which are analogs of known compounds or analogs of compounds that have been identified as "hits" or "leads" in other drug discovery screens, some of which are derived from natural products, and some of which arise from non-directed synthetic organic chemistry. Natural product libraries are collections of microorganisms, animals, plants, or marine organisms that are used to create mixtures for screening by: (1) fermentation and extraction of broths from soil, plant, or marine microorganisms or (2) extraction of plants or marine organisms. Natural product libraries include polyketides, non-ribosomal peptides, and variants (non-naturally occurring) thereof. For a review, see *Science* 282:63-68 (1998). Combinatorial libraries are composed of large numbers of peptides, oligonucleotides, or organic compounds as a mixture. These libraries are relatively easy to prepare by traditional automated synthesis methods, PCR, cloning, or proprietary synthetic methods. Of particular interest are non-peptide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, *Curr. Opin. Biotechnol.* 8:701-707 (1997). Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to modulate activity.

[0148] T1R3 binding partners that stimulate T1R3 activity are useful as agonists in disease states or conditions characterized by insufficient T1R3 signaling (*e.g.*, as a result of insufficient activity of a T1R3 ligand). T1R3 binding partners that block ligand-mediated T1R3 signaling are useful as T1R3 antagonists to treat disease states or conditions characterized by excessive T1R3 signaling. Thus, in another aspect, the invention provides methods for treating a disease or abnormal condition by administering to a patient in need of such treatment a substance that

modulates the activity or expression of a polypeptide having a sequence of SEQ ID NO:2, or exhibiting substantially the same biological activity as a polypeptide having a sequence of SEQ ID NO:2.

[0149] In addition T1R3 modulators in general, as well as T1R3 polynucleotides and polypeptides, are useful in diagnostic assays for such diseases or conditions.

Mimetics

[0150] Mimetics or mimics of compounds identified herein (sterically similar compounds formulated to mimic the key portions of the structure) may be designed for pharmaceutical use. Mimetics may be used in the same manner as the compounds identified by the present invention that modulate the T1R3 receptor and hence are also functional equivalents. The generation of a structural-functional equivalent may be achieved by the techniques of modeling and chemical design known to those of skill in the art. It will be understood that all such sterically similar constructs fall within the scope of the present invention.

[0151] The design of mimetics to a known pharmaceutically active compound is a known approach to the development of pharmaceuticals based on a "lead" compound. This is desirable where, for example, the active compound is difficult or expensive to synthesize, or where it is unsuitable for a particular method of administration, *e.g.*, some peptides may be unsuitable active agents for oral compositions as they tend to be quickly degraded by proteases in the alimentary canal.

[0152] There are several steps commonly taken in the design of a mimetic. First, the particular parts of the compound that are critical and/or important in determining its T1R3-modulating properties are determined. In the case of a polypeptide, this can be done by systematically varying the amino acid residues in the peptide, *e.g.* by substituting each residue in turn. Alanine scans of peptides are commonly used to refine such peptide motifs.

[0153] Once the active region of the compound has been identified, its structure is modeled according to its physical properties, *e.g.* stereochemistry, bonding, size, and/or charge, using data from a range of sources, such as, but not limited to, spectroscopic techniques, X-ray diffraction data, and NMR. Computational analysis, similarity mapping (which models the charge and/or volume of the active region, rather than the bonding between atoms), and other techniques known to those of skill in the art can be used in this modeling process.

[0154] In a variant of this approach, the three-dimensional structure of the compound that modulates a T1R3 receptor and the active region of the T1R3 receptor are modeled. This can be especially useful where either or both of these compounds change conformation upon binding. Knowledge of the structure of the ligand-binding domain (for example, residues 1-571 of SEQ ID NO:2) of the receptor also allows the design of high potency ligands and/or modulators.

[0155] A template molecule is then selected onto which chemical groups that mimic the T1R3 modulator can be grafted. The template molecule and the chemical groups grafted onto it can conveniently be selected so that the mimetic is easy to synthesize, is pharmacologically acceptable, and does not degrade *in vivo*, while retaining the biological activity of the lead compound. Alternatively, where the mimetic is peptide-based, further stability can be achieved by cyclizing the peptide, thereby increasing its rigidity. The mimetic or mimetics found by this approach can then be screened by the methods of the present invention to see whether they have the ability to modulate the T1R3 receptor. Further optimization or modification can then be performed to arrive at one or more final mimetics for *in vivo* or clinical testing.

Compositions of binding and/or modulating compounds

[0156] Following identification of a compound that binds and/or or modulates a T1R3 receptor, the compound may be manufactured and/or used in preparation of compositions including, but not limited to, foods, drinks, and pharmaceutical compositions. The compositions are provided or administered to patients, including, but not limited to, avians, felines, canines, bovines, ovines, porcines, equines, rodents, simians, and humans.

[0157] Thus, the present invention extends, in various aspects, not only to compounds identified in accordance with the methods disclosed herein but also foods, drinks, pharmaceutical compositions, drugs, or other compositions comprising such a compound; methods comprising administration of such a composition to a patient, *e.g.* for treatment (which includes prophylactic treatment) of a T1R3-associated disorder (*e.g.*, obesity, diabetes); uses of such a compound in the manufacture of a composition for administration to a patient; and methods of making a composition comprising admixing such a compound with a pharmaceutically acceptable excipient, vehicle or carrier, and optionally other ingredients.

[0158] The compositions of the invention comprise a taste-modifying amount of at least one or more binding or modulating compounds. A "taste-modifying amount" is a quantity sufficient to increase or decrease the perception of a taste stimulus by a given mammal. The food and drink

compositions of the invention are formulated by the addition of a binding or modulating compound to a food or drink of the mammal. Such compositions may be individualized or breed-specific. For example, feline veterinary specialty diets may thus be made more palatable.

[0159] The pharmaceutical compositions of the invention comprise a therapeutically effective amount of a compound identified according to the methods disclosed herein, or a pharmaceutically acceptable salt thereof, and a pharmaceutically acceptable carrier or excipient.

[0160] The compounds of the invention can be formulated as neutral or salt forms. Pharmaceutically acceptable salts include those formed with free amino groups such as those derived from hydrochloric, phosphoric, acetic, oxalic, tartaric acids, *etc.*, and those formed with free carboxyl groups such as those derived from sodium, potassium, ammonium, calcium, ferric hydroxides, isopropylamine, triethylamine, 2-ethylamino ethanol, histidine, procaine, *etc.*

[0161] Pharmaceutically acceptable carriers include but are not limited to saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. The carrier and composition can be sterile. The formulation should suit the mode of administration.

[0162] The composition, if desired, can also contain minor amounts of wetting or emulsifying agents or pH buffering agents. The composition can be a liquid solution, suspension, emulsion, tablet, pill, capsule, sustained release formulation, or powder. The composition can be formulated as a suppository, with traditional binders and carriers such as triglycerides. Oral formulations can include standard carriers such as pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, *etc.*

[0163] The pharmaceutical compositions of the invention may further comprise a secondary compound for the treatment of a disorder unrelated to the T1R3 receptor, such as an antibiotic or other therapeutic agent, to improve the palatability of the pharmaceutical composition, thereby improving the ease of administration.

[0164] In one embodiment, the composition is formulated in accordance with routine procedures as a pharmaceutical composition adapted for oral (*e.g.*, tablets, granules, syrups) or non-oral (*e.g.*, ointments, injections) administration to the subject. Various delivery systems are known and can be used to administer a compound that modulates a T1R3 receptor, *e.g.*, encapsulation in liposomes, microparticles, microcapsules, expression by recombinant cells, receptor-mediated endocytosis, construction of a therapeutic nucleic acid as part of a retroviral or

other vector, *etc.* Methods of introduction include but are not limited to intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal, topical, and oral routes.

[0165] The compounds of the invention may be administered by any convenient route, for example by infusion or bolus injection, by absorption through epithelial or mucocutaneous linings (*e.g.,* oral mucosa, rectal and intestinal mucosa, *etc.*), and may be administered together with other biologically active agents, for example in HAART therapy. Administration can be systemic or local. In addition, it may be desirable to introduce the pharmaceutical compositions of the invention into the central nervous system by any suitable route, including intraventricular and intrathecal injection; intraventricular injection may be facilitated by an intraventricular catheter, for example, attached to a reservoir, such as an Ommaya reservoir.

[0166] In a specific embodiment, it may be desirable to administer the pharmaceutical compositions of the invention locally to the area in need of treatment; this may be achieved by, for example, and not by way of limitation, local infusion during surgery; topical application, *e.g.*, in conjunction with a wound dressing after surgery; by injection; by means of a catheter; by means of a suppository; or by means of an implant, said implant being of a porous, non-porous, or gelatinous material, including membranes, such as sialastic membranes, or fibers.

[0167] The composition can be administered in unit dosage form and may be prepared by any of the methods well known in the pharmaceutical art, for example, as described in REMINGTON'S PHARMACEUTICAL SCIENCES (Mack Publishing Co., Easton, PA). The amount of the compound of the invention that modulates a T1R3 receptor that is effective in the treatment of a particular disorder or condition will depend on factors including but not limited to the chemical characteristics of the compounds employed, the route of administration, the age, body weight, and symptoms of a patient, the nature of the disorder or condition, and can be determined by standard clinical techniques. Typically therapy is initiated at low levels of the compound and is increased until the desired therapeutic effect is achieved. In addition, *in vitro* assays may optionally be employed to help identify optimal dosage ranges. Suitable dosage ranges for intravenous administration are preferably generally about 20-500 micrograms of active compound per kilogram body weight. Suitable dosage ranges for intranasal administration are preferably generally about 0.01 pg/kg body weight to 1 mg/kg body weight. Suppositories preferably generally contain active ingredient in the range of 0.5% to 10% by weight; oral formulations preferably may contain 10% to 95% active ingredient. Effective doses may be extrapolated from dose-response curves derived from *in vitro* or animal model test systems.

[0168] Typically, compositions for intravenous administration are solutions in sterile isotonic aqueous buffer. Where necessary, the composition may also include a solubilizing agent and a local anesthetic such as lidocaine to ease pain at the site of the injection. Generally, the ingredients are supplied either separately or mixed together in unit dosage form, for example, as a dry-lyophilized powder or water-free concentrate in a hermetically sealed container such as an ampoule or sachette indicating the quantity of active agent. Where the composition is to be administered by infusion, it can be dispensed with an infusion bottle containing sterile pharmaceutical grade water or saline.

[0169] Where the composition is administered by injection, an ampoule of sterile water for injection or saline can be provided so that the ingredients may be mixed prior to administration.

Treatment Methods

[0170] The invention provides methods of treatment of T1R3 receptor-associated disorders by administering to a subject or patient an effective amount of a compound that modulates the T1R3 receptor. In some aspects of the invention, the compounds or pharmaceutical compositions of the invention are administered to a patient having an increased risk of or having a disorder associated with the T1R3 receptor. The patient may be, for example, avian, feline, canine, bovine, ovine, porcine, equine, rodent, simian, or human.

Kits

[0171] A kit of the invention comprises a carrier means being compartmentalized to receive in close confinement one or more container means such as vials, tubes, and the like, each of the container means comprising an element to be used in the methods of the invention. For example, one of the container means may comprise the a polynucleotide encoding a T1R3 receptor of the invention, a T1R3 receptor of the invention, or an antibody thereto. The kit may also have one or more conventional kit components, including, but not limited to, instructions, test tubes, Eppendorf™ tubes, labels, reagents helpful for quantification of marker gene expression, *etc.*

EXAMPLES

[0172] The following examples are meant to be illustrative of the present invention and are not intended to limit the scope thereof.

Cloning and Characterization of the Feline T1R3 receptor

[0173] The discovery of feline taste receptor, T1R3, was achieved by using a molecular strategy termed "overgo" (Thomas, *et al.*, *Genome Res.*, 12:1277-1285 (2002); Vollrath, D., *DNA markers for physical mapping* In GENOME ANALYSIS: A LABORATORY MANUAL, Vol. 4, ed. B. Birren, *et al.*, pp. 187–215, 1999). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.). This strategy involves the use of the shortest DNA probes among the many kinds of probes used in bacterial artificial chromosome (BAC) library screening. These probes are comprised of two DNA sequences (*e.g.*, 22mers or 24mers) with a complementary 8 base overlap. They can be designed by computer program (genome.wustl.edu/tools/?overgo=1) and are readily synthesized.

[0174] Overgo probes were designed from conserved regions of the chromosome 1 marker, "disheveled 1" (DVL1) and the G protein-coupled receptor, T1R3, by aligning DVL1 and T1R3 genomic sequences from many different species. The overlapping sequences of the seven DVL1 overgo probes used in the present invention were as follows:

```
catOV1a ACTTTGAGAACATGAGTAATGACG (SEQ ID NO:21)
catOV1b AGTACCCGGACTGCGTCGTCATTA (SEQ ID NO:22)

catOV2a CACTAGGGTCATCCTTGCTTTCAG (SEQ ID NO:23)
catOV2b AGTCAGGGTGATGGGCCTGAAAGC (SEQ ID NO:24)

Ov8-OVa ATGTGGTGGACTGGCTGTACCATC (SEQ ID NO:25)
Ov8-OVb TTGAAGCCCTCCACGTGATGGTAC (SEQ ID NO:26)

Ov9a CACACGGTGAACAAGATCACCTTC (SEQ ID NO:27)
Ov9b AGTAGCACTGCTCGGAGAAGGTGA (SEQ ID NO:28)

Ov10a  ATCTACCACATGGACGAGGAGGAG (SEQ ID NO:29)
Ov10b  TGACCAGGTACGGCGTCTCCTCCT (SEQ ID NO:30)

Ov11a  AGCGCGTCACGCTGGCCGACTTCA (SEQ ID NO:31)
Ov11b  TTGCTGAGCACGTTCTTGAAGTCG (SEQ ID NO:32)

Ov12a  CACGCCTACAAATTCTTCTTTAAG (SEQ ID NO:33)
Ov12b  AGTCCTGGTCCATGGACTTAAAGA (SEQ ID NO:34).
```

The overlapping sequences of the twelve T1R3 overgo probes used in the present invention were as follows:

```
t1r3-OV1a CTTCCACTCCTGCTGCTACGACTG (SEQ ID NO:35)
t1r3-OV1b TGCCTCGCAGTCCACGCAGTCGTA (SEQ ID NO:36)

t1r3-OV2a AGGTGCGCCGCGTCAAGGGCTTCC (SEQ ID NO:37)
t1r3-OV2b TCGTAGCAGCAGGAGTGGAAGCCC (SEQ ID NO:38)

t1r3-OV3a GTTCCTGGCATGGGGGGAGCCGGC (SEQ ID NO:39)
```

```
tlr3-OV3b GAGCAGCACAAGCACAGCCGGCTC (SEQ ID NO:40)

tlr3-OV4a ACAGCCCACTAGTTCAGGCCGCAG (SEQ ID NO:41)
tlr3-OV4b CAGGCCCGGGGTCCCCCTGCGGCC (SEQ ID NO:42)

tlr3-OV5a CCCACTGGTTCAGGCCTCGGGGGG (SEQ ID NO:43)
tlr3-OV5b AAAGCAGGCCAGGGGCCCCCCCGA (SEQ ID NO:44)

tlr3-OV6a AGGCGCTGGTGCACTGCCGCACAC (SEQ ID NO:45)
tlr3-OV6b AAGCTGACCCAGGAGCGTGTGCGG (SEQ ID NO:46)

tlr3-OV7a ACAGAGGCACTGGTGCACTGCCGC (SEQ ID NO:47)
tlr3-OV7b TGATCCAGGAGTGCACGCGGCAGT (SEQ ID NO:48)

tlr3-OV8a ACCAATGCCACGCTGGCCTTTCTC (SEQ ID NO:49)
tlr3-OV8b AAGTGCCCAGGAAGCAGAGAAAGG (SEQ ID NO:50)

tlr3-OV9a TGGTACATGCTGCCAATGCCACGC (SEQ ID NO:51)
tlr3-OV9b AAGCAGAGGAAAGCCAGCGTGGCA (SEQ ID NO:52)

tlr3-OV10a TACAACCGTGCCCGTGGCCTCACC (SEQ ID NO:53)
tlr3-OV10b AGGCCAGCATGGCGAAGGTGAGGC (SEQ ID NO:54)

tlr3-OV11a TCATCACCTGGGTCTCCTTTGTGC (SEQ ID NO:55)
tlr3-OV11b ACATTGGCCAGGAGGGGCACAAAG (SEQ ID NO:56)

tlr3-OV12a TGCAGATGGGTGCCCTCCTGCTCT (SEQ ID NO:57)
tlr3-OV12b AGGATGCCCAGCACACAGAGCAGG (SEQ ID NO:58).
```

The single-stranded overhangs were filled in with ^{32}P labeled dATP and dCTP, and the overgo probes hybridized with BAC libraries.

[0175] The overgo strategy is considered to be more versatile than a PCR-based strategy by those skilled in the art of comparative physical mapping for the following reasons: (1) overgo probes are short (*e.g.*, 36mers or 40mers), making the probability of good alignment from among many species more favorable; (2) overgo probes are more specific to the target genes compared with traditional cDNA and genomic DNA probes used by PCR; and (3) although overgo probes are short, they are not as restricted as traditional PCR probes, which cannot tolerate even a few mismatches, because they can be used in hybridization approaches with BACs or other libraries.

[0176] **Screening a feline genomic BAC library.** Seven DVL1 overgo probes (SEQ ID NOS:21-34) were used in screening a feline genomic BAC library. Probes were radioactively labeled by the random hexa-nucleotide method (Feinberg & Vogelstein, *Analytical Biochemistry*, **132**:6-13 (1983)). Hybridization and washing of membranes followed standard protocols (Church & Gilbert, *PNAS U.S.A.*, 81:1991-1995 (1984)). Thirty-nine positive BAC clones were

identified. Several BAC ends were sequenced. One clone containing homologous sequence to human chromosome 1p36, BAC 552J19, was identified using bioinformatics tools.

[0177] **Production of a shotgun library for BAC 552J19 and identification of a single clone containing feline T1R3.** BAC DNA from 552J19 was prepared by using Qiagen Large Construct Kit. DNA was then digested by the restriction enzyme *Sau3A1* and subcloned into pGEM+3Z (Promega) vector. After transformants were arrayed to a nylon membrane, two separate hybridizations were performed using seven DVL1 and twelve T1R3 overgo probes (SEQ ID NOS:35-58). Two clones positive for DVL1 and four clones positive for T1R3 were found. These clones were confirmed by sequencing. Because DVL1 is the neighboring gene of T1R3 in human and mouse, it is likely this also is the case in cat; therefore, the DVL1 positive clones verified that the BAC 552J19 is the correct BAC, that is, it is the one containing feline T1R3.

Results

[0178] More than 3 kb of genomic sequences containing the open reading frame for domestic cat taste receptor, T1R3, were obtained. Figure 1 shows the multiple sequence alignments of the known nucleotide sequences for the T1R receptors human (T1R1, SEQ ID NO:8; T1R2, SEQ ID NO:5; T1R3, SEQ ID NO:11), mouse (T1R1, SEQ ID NO:6; T1R2, SEQ ID NO:3; T1R3, SEQ ID NO:9), and rat (T1R1, SEQ ID NO:7; T1R2, SEQ ID NO:4; T1R3, SEQ ID NO:10), along with the newly discovered and novel nucleotide sequence for the T1R3 taste receptor from domestic cat (SEQ ID NO:1).

[0179] Figure 2 shows the deduced amino acid sequence of the domestic cat taste receptor, T1R3 (SEQ ID NO:2), aligned with the amino acid sequences of the T1R receptor family human (T1R1, SEQ ID NO:17; T1R2, SEQ ID NO:20; T1R3, SEQ ID NO:12), rat (T1R1, SEQ ID NO:16; T1R2, SEQ ID NO:19; T1R3, SEQ ID NO:14), and mouse (T1R1, SEQ ID NO:15; T1R2, SEQ ID NO:18; T1R3, SEQ ID NO:13). The deduced cat sequence predicts four additional amino acids at positions 11 – 14 relative to the homologous T1R3 receptors of mouse, human, and rat. The deduced sequence for cat reveals a threonine in position 64, a position equivalent to amino acid 60 in mouse, and a leucine at position 59, a position equivalent to position 55 in mouse. In mouse, amino acid substitutions of a threonine at position 60 and an alanine at position 55, both positions located within the putative extracellular N-terminal domain of the polypeptide, are present in strains of mice demonstrating low preference for the sweet stimulus saccharin (Bachmanov *et al.*, *Chem. Senses*, 26:925-933 (2001)). Leucine is a

conservative substitution for alanine. Accordingly, the amino acid sequence differences of cat and mouse T1R3 receptor may account for functional differences that lead to different taste preferences between the two species. For example, the amino acid substitutions may explain the cat's inability to taste many compounds that have a sweet taste to mice and humans.

[0180] Table 4 shows the percent homology among the members of the T1R family in relation to the newly discovered cat T1R3 taste receptor. The portion of Table 4 to the left of the diagonal (in bold type) shows the percent homology based on the open reading frame of the nucleotide sequences obtained from Figure 1 for the T1R family among human, cat, rat, and mouse. The upper portion to the right of the diagonal (in italic type) shows the percent homology of the T1R members based on the amino acid sequences of Figure 2. Cat T1R3 shows 79% nucleotide sequence homology with human T1R3, 75% with rat T1R3 and 74% with mouse T1R3. At the amino acid level, cat T1R3 shows 73% homology with human T1R3, 72% with rat, and 72% with mouse. Cat T1R3 shows generally low homology with the other known members of the T1R family, T1R1 and T1R2, from human, rat, and mouse. The same range of relatively low homology is present among the human, rat, and mouse T1R3 and the T1R1 and T1R2 receptors from the same species.

Table 4. Percent Homology Among Diverse Species for T1Rs

Species	Mouse T1R1	Mouse T1R2	Mouse T1R3	Rat T1R1	Rat T1R2	Rat T1R3	Human T1R1	Human T1R2	Human T1R3	Cat T1R3
Mouse T1R1		*36*	*30*	*90*	*36*	*30*	*73*	*37*	*30*	*30*
Mouse T1R2	**55**		*28*	*36*	*91*	*28*	*34*	*69*	*28*	*28*
Mouse T1R3	**33**	**15**		*31*	*28*	*92*	*30*	*27*	*72*	*72*
Rat T1R1	**91**	**55**	**33**		*37*	*31*	*73*	*37*	*31*	*31*
Rat T1R2	**55**	**91**	**15**	**57**		*28*	*34*	*71*	*29*	*28*
Rat T1R3	**33**	**21**	**93**	**32**	**15**		*31*	*27*	*73*	*72*
Human T1R1	**79**	**56**	**35**	**79**	**56**	**35**		*35*	*31*	*31*
Human T1R2	**57**	**78**	**17**	**56**	**78**	**17**	**57**		*28*	*28*
Human T1R3	**41**	**39**	**73**	**39**	**36**	**75**	**40**	**38**		*73*
Cat T1R3	**33**	**34**	**74**	**36**	**36**	**75**	**53**	**39**	**79**	

Note: Upper right cells (*italics*) contain deduced amino acid homology; lower left cells (**bold**) contain nucleotide homology.

[0181] The rat and mouse have closely related T1R receptors, while the T1R3 of human and cat diverge from these two, as illustrated in the phylogenetic tree of Figure 3. Interestingly, the types of sweet compounds to which the rat and mouse respond are very similar, whereas those that stimulate the human and those that stimulate the cat are much different from those for rat and mouse, and whereas the compounds that stimulate the cat and human receptors also are very different.

[0182] The feline T1R3 receptor is a seven transmembrane receptor similar in structure to other known members of the T1R family of receptors (Figure 4). The structure of the feline T1R3 receptor was generated through use of a protein modeling program available at <www.ebi.ac.uk/~moeller/transmembrane.html>.

What is Claimed:

1. An isolated and purified polynucleotide encoding a T1R3 receptor comprising:

a) the nucleotide sequence of SEQ ID NO:1,

b) a fragment of at least about 42 contiguous nucleotides of SEQ ID NO:1 encoding a polypeptide having substantially the same biological activity as a polypeptide encoded by the nucleotide sequence of SEQ ID NO:1,

c) a variant of the polynucleotide of SEQ ID NO:1 having at least 80% homology to the polynucleotide of SEQ ID NO:1 and encoding a polypeptide having substantially the same biological activity as a polypeptide encoded by the nucleotide sequence of SEQ ID NO:1,

d) a variant of the polynucleotide of SEQ ID NO:1 having at least 80% homology to the polynucleotide of SEQ ID NO:1 and encoding a polypeptide conferring modified taste perception to one or more taste stimuli relative to a polypeptide encoded by the polynucleotide of SEQ ID NO:1,

e) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO:2,

f) a nucleotide sequence substantially complementary to the nucleotide sequence of SEQ ID NO:1, or

g) a nucleotide sequence that hybridizes to the complement of the polynucleotide having SEQ ID NO:1 under high stringency conditions.

2. The polynucleotide of claim 1, wherein said polynucleotide is DNA.

3. The polynucleotide of claim 1, wherein said polynucleotide is RNA.

4. The polynucleotide of claim 1 comprising a variant of the polynucleotide of SEQ ID NO:1 encoding an amino acid sequence of SEQ ID NO:2 having a nonconserved amino acid substitution at residue 59 or residue 64.

5. The polynucleotide of claim 1 comprising a fragment of the polynucleotide of SEQ ID NO:1 wherein said fragment comprises a nucleotide sequence encoding an extracellular domain of the polypeptide of SEQ ID NO:2, a transmembrane domain of the polypeptide of SEQ ID NO:2, or an intracellular domain of the polypeptide of SEQ ID NO:2.

6. An expression vector comprising the polynucleotide of claim 1 operably linked to a promoter.

7. A host cell comprising the expression vector of claim 6.

8. The host cell of claim 7 wherein said cell is mammalian.

9. The host cell of claim 8 wherein said cell is feline.

10. A cell culture comprising at least one cell of claim 6.

11. An isolated and purified T1R3 receptor polypeptide encoded by a polynucleotide of claim 1.

12. The polypeptide of claim 11 wherein said polypeptide comprises the amino acid sequence of SEQ ID NO:2, a fragment of at least 30 contiguous amino acids of SEQ ID NO:2, or a variant thereof having substantially the same biological activity as the polypeptide of SEQ ID NO:2.

13. The polypeptide of claim 11 wherein said polypeptide comprises an amino acid sequence having at least one sequence variation of SEQ ID NO:2 wherein said variation confers modified taste perception to one or more taste stimuli relative to a polypeptide of SEQ ID NO:2.

14. An isolated and purified T1R3 receptor polypeptide comprising the amino acid sequence of SEQ ID NO:2.

15. The polypeptide of claim 12, wherein said polypeptide comprises a feline T1R3 receptor.

16. A kit for the detection of a polynucleotide encoding a feline T1R3 receptor comprising a polynucleotide that specifically hybridizes to a polynucleotide encoding a polypeptide having an amino acid sequence of SEQ ID NO:2 and instructions relating to detection of said polynucleotide.

17. An isolated and purified antibody that immunoreacts specifically with at least one epitope of a polypeptide of one of claims 11, 12, 13, 14, or 15.

18. A kit for the detection of a polypeptide encoding a feline T1R3 receptor comprising the antibody of claim 17 and instructions relating to detection.

19. A method of producing a feline T1R3 receptor comprising culturing the host cell of claim 7 and recovering said receptor from said host cell.

20. The feline T1R3 receptor produced according to the method of claim 19.

21. A method for identifying compounds that interact with a feline T1R3 receptor comprising:

contacting a feline T1R3 receptor of claim 11, 12, 13, 14, or 15 with a test compound, and

detecting interaction between said receptor and said compound.

22. The method of claim 21, wherein said receptor is bound to a solid support.

23. The method of claim 22, wherein said solid support is formulated into a feline-specific electronic tongue.

24. A method for identifying an agonist of a feline T1R3 receptor comprising:

expressing an expression vector of claim 6 in the presence of a test compound, and

detecting an increase in biological activity of a polypeptide produced by said expression step in the presence of said compound relative to biological activity of said polypeptide in the absence of said compound.

25. A method for identifying an agonist of a feline T1R3 receptor comprising:

contacting a polypeptide of claim 11, 12, 13, 14, or 15 with a test compound, and

detecting an increase in biological activity of said polypeptide in the presence of said compound relative to biological activity of said polypeptide in the absence of said compound.

26. A method for identifying an antagonist of a feline T1R3 receptor comprising:

expressing an expression vector of claim 6 in the presence of a test compound, and

detecting a decrease in biological activity of a polypeptide produced by said expression step in the presence of said compound relative to biological activity of said polypeptide in the absence of said compound.

27. A method for identifying an antagonist of a feline T1R3 receptor comprising:

contacting the polypeptide of claim 11, 12, 13, 14, or 15 with a test compound, and

detecting a decrease in biological activity of said polypeptide in the presence of said compound relative to biological activity of said polypeptide in the absence of said compound.

28. A compound for modifying the taste perception of a mammal identified according to the method of claim 24.

29. A compound for modifying the taste perception of a mammal identified according to the method of claim 25.

30. A compound for modifying the taste perception of a mammal identified according to the method of claim 26.

31. A compound for modifying the taste perception of a mammal identified according to the method of claim 27.

32. A composition for modifying the taste perception of a mammal comprising at least one compound identified by the method of claim 24.

33. A composition for modifying the taste perception of a mammal comprising at least one compound identified by the method of claim 25.

34. A composition for modifying the taste perception of a mammal comprising at least one compound identified by the method of claim 26.

35. A composition for modifying the taste perception of a mammal comprising at least one compound identified by the method of claim 27.

36. The composition of claim 32, wherein said composition comprises a feline food, drink, or pharmaceutical composition.

37. The composition of claim 33, wherein said composition comprises a feline food, drink, or pharmaceutical composition.

38. The composition of claim 34, wherein said composition comprises a feline food, drink, or pharmaceutical composition.

39. The composition of claim 35, wherein said composition comprises a feline food, drink, or pharmaceutical composition.

40. The composition of claim 36, wherein said composition is breed-specific.

41. The composition of claim 37, wherein said composition is breed-specific.

42. The composition of claim 38, wherein said composition is breed-specific.

43. The composition of claim 39, wherein said composition is breed-specific.

44. A method for modifying the taste perception of a mammal comprising administering to said mammal a compound identified according to the method of claim 24.

45. A method for modifying the taste perception of a mammal comprising administering to said mammal a compound identified according to the method of claim 25.

46. A method for modifying the taste perception of a mammal comprising administering to said mammal a compound identified according to the method of claim 26.

47. A method for modifying the taste perception of a mammal comprising administering to said mammal a compound identified according to the method of claim 27.

48. A composition for modifying the taste perception of a mammal comprising at least one polynucleotide of claim 1 and a pharmaceutically acceptable excipient.

49. A method for modifying the taste perception of a mammal comprising administering to said mammal at least one polynucleotide of claim 1.

50. A method for modifying the taste perception of a mammal comprising administering to said mammal at least one polypeptide of claim 11.

51. A transgenic animal comprising a polynucleotide of claim 1.

52. The method of claim 25 wherein said polypeptide is bound to a solid support.

53. The method of claim 52 wherein said solid support is formulated into a feline-specific electronic tongue.

54. The method of claim 27 wherein said polypeptide is bound to a solid support.

55. The method of claim 54 wherein said solid support is formulated into a feline-specific electronic tongue.

56. A method of identifying a feline T1R3 receptor variant that confers modified taste perception comprising expressing a variant of the polynucleotide of SEQ ID NO:1 having at least 80% homology to the polynucleotide of SEQ ID NO:1 and detecting an increase or a decrease in the biological activity of the polypeptide encoded by the variant relative to the biological activity of the polypeptide encoded by SEQ ID NO:1.

ABSTRACT

The present invention relates to the discovery of a gene of the domestic cat (*Felis catus*) associated with taste perception. The invention provides, *inter alia*, the nucleotide sequence of the feline T1R3 receptor gene, the amino acid sequence of the polypeptide encoded thereby, and antibodies to the polypeptide. The present invention also relates to methods for screening for compounds that modify the gene's function or activity, the compounds identified by such screens, and mimetics of the identified compounds. The invention further provides methods for modifying the taste preferences, ingestive responses, or general behavior of a mammal by administering compounds that affect the function or activity of the gene.

Figure 1A
CLUSTAL W (1.82) multiple nucle tide sequence alignment of T1Rs:

```
mouseTas1r2     ATGGGACCCCAGGCGAG------GACACTCC-ATTTGCTGTTTCTC-CTGCTGCATGCTC 52
ratTas1r2       ATGGGTCCCCAGGCAAG------GACACTCT-GCTTGCTGTCTCTC-CTGCTGCATGTTC 52
humanTAS1R2     ATGGGGCCCAGGGCAAA------GACCATCT-GCTCCCTGTTCTTC-CTCCTATGGGTCC 52
mouseTas1r1     ATGCTTTTCTGGGCAGCTCACCTGCTGCTCA-GCCTGCAGCTGGCCGTTGCTTACTGCTG 59
ratTas1r1       ATGCTCTTCTGGGCTGCTCACCTGCTGCTCA-GCCTGCAGTTGGTC------TACTGCTG 53
humanTAS1R1     ATGCTGCTCTGCACGGCTCGCCTGGT---CG-GCCTGCAGCTTCTCATTTCCTGCTGCTG 56
mouseTas1r3     ATGCCAGCTTTGGCTAT---CATGGGTCTCA------------GCCTGGCTGCTTTCCTG 45
ratTas1r3       ATGCCGGGTTTGGCTAT---CTTGGGCCTCA------------GTCTGGCTGCTTTCCTG 45
catTas1r3       ATGCCCGGCCTCGCTCT---CCTGGGCCTCACGGCTCTCCTGGGCCTCACGGCTCTCTTG 57
humanTAS1R3     ATGCTGGGCCCTGCTGT---CCTGGGCCTCA------------GCCTCTGGGCTCTCCTG 45
                ***          *         *     *               *

mouseTas1r2     TGCCTAAGCCAGTCATGCTGGTAGGGAAC-TCC---GACTTTCACCTGGCTGGGGACTAC 108
ratTas1r2       TGCCTAAGCCAGGCAAGCTGGTAGAGAAC-TCT---GACTTCCACCTGGCCGGGGACTAC 108
humanTAS1R2     TGGCTGAGCC---------GGCTGAGAAC-TCG---GACTTCTACCTGCCTGGGGATTAC 99
mouseTas1r1     GGCTTTCAGCTGCCAAAG-GACAGAATCC-TCTCCAGGTTTCAGCCTCCCTGGGGACTTC 117
ratTas1r1       GGCTTTCAGCTGCCAAAG-GACAGAGTCC-TCTCCAGGCTTCAGCCTTCCTGGGGACTTC 111
humanTAS1R1     GGCCTTTGCCTGCCATAG-CACGGAGTCT-TCTCCTGACTTCACCCTCCCCGGAGATTAC 114
mouseTas1r3     GAGCTTGGGATGGGGGCCTCTTTGTGTCTGTCACAGCAATTCAAGGCACAAGGGGACTAC 105
ratTas1r3       GAGCTTGGGATGGGGTCCTCTTTGTGTCTGTCACAGCAATTCAAGGCACAAGGGGACTAT 105
catTas1r3       GACCACGGGGAGGGCGCAACGTCCTGCTTGTCACAGCAGCTCAGGATGCAGGGGGACTAT 117
humanTAS1R3     CACCCTGGGACGGGGGCCCCATTGTGCCTGTCACAGCAACTTAGGATGAAGGGGGACTAC 105
                                              **        *          ** ** *

mouseTas1r2     CTCCTGGGTGGCCTCTTTACCCTCCATGCCAACGTGAAGAGCGTCTCTCACCTCAGCTAC 168
ratTas1r2       CTCCTGGGTGGCCTCTTTACCCTCCATGCCAACGTGAAGAGCATCTCCCACCTCAGCTAC 168
humanTAS1R2     CTCCTGGGTGGCCTCTTCTCCCTCCATGCCAACATGAAGGGCATTGTTCACCTTAACTTC 159
mouseTas1r1     CTCCTGGCAGGCCTGTTCTCCCTCCATGCTGACTGTCTGCAGGTGAGACAC---AG--AC 172
ratTas1r1       CTCCTTGCAGGTCTGTTCTCCCTCCATGGTGACTGTCTGCAGGTGAGACAC---AG--AC 166
humanTAS1R1     CTCCTGGCAGGCCTGTTCCCTCTCCATTCTGGCTGTCTGCAGGTGAGGCAC---AG--AC 169
mouseTas1r3     ATACTGGGCGGGCTATTTCCCCTGGGCTCAACCG---AGGAGGCCACTCTC---AACCAG 159
ratTas1r3       ATATTGGGTGGACTATTTCCCCTGGGCACAACTG---AGGAGGCCACTCTC---AACCAG 159
catTas1r3       GTGCTGGGTGGGCTCTTCCCTCTGGGCTCTGCCG---AGGGTACAGGTCTT---GGCGAC 171
humanTAS1R3     GTGCTGGGGGGGCTGTTCCCCCTGGGCGAGGCCG---AGGAGGCTGGCCTC---CGCAGC 159
                 *  * *  ** ** **  * **               *         *

mouseTas1r2     CTGCAGGTGCCCAAGTGCAATGAGTACAACA---TGAAGGTCTTGGGCTACAACCTCATG 225
ratTas1r2       CTGCAGGTGCCCAAGTGCAATGAGTTCACCA---TGAAGGTGTTGGGCTACAACCTCATG 225
humanTAS1R2     CTGCAGGTGCCCATGTGCAAGGAGTATGAAG---TGAAGGTGATAGGCTACAACCTCATG 216
mouseTas1r1     CT-CTGGTGACAAGTTGTGACAGGTCTGACAGCTTCAACGGCCATGGCTATCACCTCTTC 231
ratTas1r1       CT-CTGGTGACAAGTTGTGACAGGCCCGACAGCTTCAACGGCCATGGCTACCACCTCTTC 225
humanTAS1R1     CC-GAGGTGACCCTGTGTGACAGGTCTTGTAGCTTCAATGAGCATGGCTACCACCTCTTC 228
mouseTas1r3     AGAACACAACCCAACAGCATCCCGTGCAACAGGTTCTCACCCCTTGGTTTGTTCCTGGCC 219
ratTas1r3       AGAACACAGCCCAACGGCATCCTATGTACCAGGTTCTCGCCCCTTGGTTTGTTCCTGGCC 219
catTas1r3       GGGCTGCAGCCCAATGCCACCGTGTGCACCAGGTTCTCGTCTCTGGGCCTGCTCTGGGCG 231
humanTAS1R3     CGGACACGGCCCAGCAGCCCTGTGTGCACCAGGTTCTCCTCAAACGGCCTGCTCTGGGCA 219
                          *                       *          **      *

mouseTas1r2     CAGGCCATGCGATTCGCCGTGGAGGAAATCAACAACTGTAGCTCTCTGCTGCCCGGCGTG 285
ratTas1r2       CAGGCCATGCGTTTCGCTGTGGAGGAGATCAACAACTGTAGCTCCCTGCTACCCGGCGTG 285
humanTAS1R2     CAGGCCATGCGCTTCGCGGTGGAGGAGATCAACAATGACAGCAGCCTGCTGCCTGGTGTG 276
mouseTas1r1     CAAGCCATGCGGTTCACCGTTGAGGAGATAAACAACTCCACAGCTCTGCTTCCCAACATC 291
ratTas1r1       CAAGCCATGCGGTTCACTGTTGAGGAGATAAACAACTCCTCGGCCCTGCTTCCCAACATC 285
humanTAS1R1     CAGGCTATGCGGCTTGGGGTTGAGGAGATAAACAACTCCACGGCCCTGCTGCCCAACATC 288
mouseTas1r3     ATGGCTATGAAGATGGCTGTGGAGGAGATCAACAATGGATCTGCCTTGCTCCCTGGGCTG 279
ratTas1r3       ATGGCTATGAAGATGGCTGTAGAGGAGATCAACAATGGATCTGCCTTGCTCCCTGGGCTG 279
catTas1r3       CTGGCCGTGAAGATGGCGGTGGAGGAGATCAACAACGGGTCGGCCCTGCTGCCCGGGCTG 291
humanTAS1R3     CTGGCCATGAAAATGGCCGTGGAGGAGATCAACAACAAGTCGGATCTGCTGCCCGGGCTG 279
                   **  **    *    ** ***** ** *****           **** **     *
```

Figure 1B

```
mouseTas1r2      CTGCTCGGCTACGAGATGGTGGATGTCTGCTACCTCTCC---AACAATATCCAGCCTGGG 342
ratTas1r2        CTGCTCGGCTACGAGATGGTGGATGTCTGTTACCTCTCC---AACAATATCCACCCTGGG 342
humanTAS1R2      CTGCTGGGCTATGAGATCGTGGATGTGTGCTACATCTCC---AACAATGTCCAGCCGGTG 333
mouseTas1r1      ACCCTGGGGTATGAACTGTATGACGTGTGCTCAGAGTCT---TCCAATGTCTATGCCACC 348
ratTas1r1        ACCCTGGGGTATGAGCTGTACGACGTGTGCTCAGAATCT---GCCAATGTGTATGCCACC 342
humanTAS1R1      ACCCTGGGGTACCAGCTGTATGATGTGTGTTCTGACTCT---GCCAATGTGTATGCCACG 345
mouseTas1r3      CGGCTGGGCTATGACCTATTTGACACATGCTCCGAGCCAGTGGTCACCATGAAATCCAGT 339
ratTas1r3        CGACTGGGCTATGACCTGTTTGACACATGCTCAGAGCCAGTGGTCACCATGAAGCCCAGC 339
catTas1r3        CACCTGGGCTATGACCTCTTTGACACGTGTTCAGAGCCCATGGTGGCCATGAAGCCCAGC 351
humanTAS1R3      CGCCTGGGCTACGACCTCTTTGATACGTGCTCGGAGCCTGTGGTGGCCATGAAGCCCAGC 339
                    ** ** **  *  *    **    ** *      *           *  *  *

mouseTas1r2      CTCTACTTCCTGTC---ACAGATAGATGACTTCCTGCCCATCCTCAAAGACTACAGCCAG 399
ratTas1r2        CTCTACTTCCTGGC---ACAGGACGACGACCTCCTGCCCATCCTCAAAGACTACAGCCAG 399
humanTAS1R2      CTCTACTTCCTGGC---ACACGAGGACAACCTCCTTCCCATCCAAGAGGACTACAGTAAC 390
mouseTas1r1      CTGAGGGTGCTCGCCCAGCAAGGGACAGGCCACCTAGAGATGCAGAGAGATCTTCGCAAC 408
ratTas1r1        CTGAGGGTGCTTGCCCTGCAAGGGCCCCGCCACATAGAGATACAGAAAGACCTTCGCAAC 402
humanTAS1R1      CTGAGAGTGCTCTCCCTGCCAGGGCAACACCACATAGAGCTCCAAGGAGACCTTCTCCAC 405
mouseTas1r3      CTCATGTTCCTGGCCAAGGTGGGCAGTCAAAGCATTGCTGCCTACTGCAACTACACACAG 399
ratTas1r3        CTCATGTTCATGGCCAAGGTGGGAAGTCAAAGCATTGCTGCCTACTGCAACTACACACAG 399
catTas1r3        CTCGTGTTCATGGCCAAAGCAGGCAGCTGCAGCATTGCCGCCTACTGCAATTACACACAG 411
humanTAS1R3      CTCATGTTCCTGGCCAAGGCAGGCAGCCGCGACATCGCCGCCTACTGCAACTACACGCAG 399
                 **     *  *  *                  * *              *        *

mouseTas1r2      TACAGGCCCCAAGTGGTGGCCGTCATTGGCCCAGACAACTCTGAGTCCGCCATCACCGTG 459
ratTas1r2        TACATGCCCCACGTGGTGGCTGTCATTGGCCCCGACAACTCTGAGTCCGCCATTACCGTG 459
humanTAS1R2      TACATTTCCCGTGTGGTGGCTGTCATTGGCCCTGACAACTCCGAGTCTGTCATGACTGTG 450
mouseTas1r1      CACTCCTCCAAGGTGGTGGCACTCATTGGGCCTGATAACACTGACCACGCTGTCACCACT 468
ratTas1r1        CACTCCTCCAAGGTGGTGGCCTTCATCGGGCCTGACAACACTGACCACGCTGTCACTACC 462
humanTAS1R1      TATTCCCCTACGGTGCTGGCAGTGATTGGGCCTGACAGCACCAACCGTGCTGCCACCACA 465
mouseTas1r3      TACCAACCCCGTGTGCTGGCTGTCATCGGCCCCCACTCATCAGAGCTTGCCCTCATTACA 459
ratTas1r3        TACCAACCCCGTGTGCTGGCTGTCATTGGTCCCCACTCATCAGAGCTTGCCCTCATTACA 459
catTas1r3        TACCAGCCCCGCGTGCTGGCCGTCATCGGGCCCCACTCGTCTGAGCTCGCCCTCGTCACC 471
humanTAS1R3      TACCAGCCCCGTGTGCTGGCTGTCATCGGGCCCCACTCGTCAGAGCTCGCCATGGTCACC 459
                  *     *    *** ****  * ** ** **  *     *  *    *

mouseTas1r2      TCCAACATTCTCTCCTACTTCCTCGTGCCACAGGTCACATATAGCGCCATCACCGACAAG 519
ratTas1r2        TCCAACATTCTCTCTCATTTCCTCATCCCACAGATCACATACAGCGCCATCTCCGACAAG 519
humanTAS1R2      GCCAACTTCCTCTCCCTATTTCTCCTTCCACAGATCACCTACAGCGCCATCAGCGATGAG 510
mouseTas1r1      GCTGCCCTGCTGAGCCCTTTTCTGATGCCCCTGGTCAGCTATGAGGCGAGCAGCGTGATC 528
ratTas1r1        GCTGCCTTGCTGGGTCCTTTCCTGATGCCCCTGGTCAGCTATGAGGCAAGCAGCGTGGTA 522
humanTAS1R1      GCCGCCCTGCTGAGCCCTTTCCTGGTGCCCATGATTAGCTATGCGGCCAGCAGCGAGACG 525
mouseTas1r3      GGCAAGTTCTTCAGCTTCTTCCTCATGCCACAGGTCAGCTATAGTGCCAGCATGGATCGG 519
ratTas1r3        GGCAAGTTCTTCAGCTTCTTCCTCATGCCACAGGTCAGCTATAGTGCCAGCATGGATCGG 519
catTas1r3        GGCAAGTTCTTCAGCTTCTTCCTTGTGCCTCAGGTCAGCTACGGCGCCAGCACCGACCGG 531
humanTAS1R3      GGCAAGTTCTTCAGCTTCTTCCTCATGCCCCAGGTCAGCTACGGTGCTAGCATGGAGCTG 519
                        *  *       ** **  * **   * * *  **    ** * *   *

mouseTas1r2      CTGCGAGACAAGCGGCGCTTCCCTGCCATGCTGCGCACTGTGCCCAGCGCCACCCACCAC 579
ratTas1r2        CTGCGGGACAAGCGGCACTTCCCTAGCATGCTACGCACAGTGCCCAGCGCCACCCACCAC 579
humanTAS1R2      CTGCGAGACAAGGTGCGCTTCCCGGCTTTGCTGCGTACCACACCCAGCGCCGACCACCAC 570
mouseTas1r1      CTCAGTGGGAAGCGCAAGTTCCCGTCCTTCTTGCGCACCATCCCCAGCGATAAGTACCAG 588
ratTas1r1        CTCAGTGCCAAGCGCAAGTTCCCGTCTTTCCTTCGTACCGTCCCCAGTGACCGGCACCAG 582
humanTAS1R1      CTCAGCGTGAAGCGGCAGTATCCCTCTTTCCTGCGCACCATCCCCAATGACAAGTACCAG 585
mouseTas1r3      CTAAGTGACCGGGAAACGTTTCCATCCTTCTTCCGCACAGTGCCCAGTGACCGGGTGCAG 579
ratTas1r3        CTAAGTGACCGGGAAACATTTCCATCCTTCTTCCGCACAGTGCCCAGTGACCGGGTGCAG 579
catTas1r3        CTGAGCAACCGGGAGATCTTCCCGTCCTTCTTCCGCACGGTGCCCAGCGACCAGGTGCAG 591
humanTAS1R3      CTGAGCGCCCGGGAGACCTTCCCCTCCTTCTTCCGCACCGTGCCCAGCGACCGTGTGCAG 579
                 **  *      *      *  **     *  * ** **    ****  *        **

mouseTas1r2      ATCGAGGCCATGGTGCAACTGATGGTTCACTTCCAGTGGAACTGGATCGTGGTGCTGGTG 639
ratTas1r2        ATCGAGGCCATGGTGCAGCTGATGGTTCACTTCCAATGGAACTGGATTGTGGTGCTGGTG 639
humanTAS1R2      GTCGAGGCCATGGTGCAGCTGATGCTGCACTTCCGCTGGAACTGGATCATTGTGCTGGTG 630
mouseTas1r1      GTGGAAGTCATAGTGCGGCTGCTGCAGAGCTTCGGCTGGGTCTGGATCTCGCTCGTTGGC 648
ratTas1r1        GTGGAGGTCATGGTGCAGCTGCTGCAGAGTTTTGGGTGGGTGTGGATCTCGCTCATTGGC 642
humanTAS1R1      GTGGAGACCATGGTGCTGCTGCTGCAGAAGTTCGGGTGGACCTGGATCTCTCTGGTTGGC 645
mouseTas1r3      CTGCAGGCAGTTGTGACTCTGTTGCAGAACTTCAGCTGGAACTGGGTGGCCGCCTTAGGG 639
ratTas1r3        CTGCAGGCCGTTGTGACACTGTTGCAGAATTTCAGCTGGAACTGGGTGGCTGCCTTAGGT 639
catTas1r3        GTGGCGGCCATGGTGGAGCTGCTGGAGGAGCTCGGCTGGAACTGGGTGGCGGCGGTGGGT 651
humanTAS1R3      CTGACGGCCGCCGCGGAGCTGCTGCAGGAGTTCGGCTGGAACTGGGTGGCCGCCCTGGGC 639
                  *          * *   *** **       *    ***   *** *        * *
```

Docket No.: MON-3042
App No.: Not Yet Assigned Filed: Herewith
Title: TASTE RECEPTOR OF THE T1R FAMILY FROM DOMESTIC CAT
Inventors: Xia LI, Weihua LI, Danielle R. REED, Alexander Alexeyevich BACHMANOV, Joseph G. BRAND
Attorney: Felicity E. Groth Phone: (215) 568-3100
Sheet 3 of 13

Figur 1C

```
mouseTas1r2         AGCGATGACGATTATGGCCGAGAGAACAGCCACCTGCTGAGCCAGCGTCTGACCAACACT 699
ratTas1r2           AGCGACGACGATTACGGCCGCGAGAACAGCCACCTGTTGAGCCAGCGTCTGACCAAAACG 699
humanTAS1R2         AGCAGCGACACCTATGGCCGCGACAATGGCCAGCTGCTTGGCGAGCGCGTGGCCCG---G 687
mouseTas1r1         AGCTATGGTGACTACGGGCAGCTGGGCGTACAGGCGCTGGAGGAGC---TGGCCACTCCA 705
ratTas1r1           AGCTACGGTGATTACGGGCAGCTGGGTGTGCAGGCGCTGGAGGAGC---TGGCCGTGCCC 699
humanTAS1R1         AGCAGTGACGACTATGGGCAGCTAGGGGTGCAGGCACTGGAGAACC---AGGCCACTGGT 702
mouseTas1r3         AGTGATGATGACTATGGCCGGGAAGGTCTGAGCATCTTTTCTAGTC---TGGCCAATGCA 696
ratTas1r3           AGTGATGATGACTATGGCCGGGAAGGTCTGAGCATCTTTTCTGGTC---TGGCCAACTCA 696
catTas1r3           AGTGACGACGAGTATGGCCGGCAGGGCCTGAGCCTCTTCTCCGGCC---TGGCCAGCGCC 708
humanTAS1R3         AGCGACGACGAGTACGGCCGGCAGGGCCTGAGCATCTTCTCGGCCC---TGGCCGCGGCA 696
                    **    *     ** ** *                  *       *    * **

mouseTas1r2         GGCGATATCTGCATTGCCTTCCAGGAGGTTCTGCCTGTACCAGAACCCAACCAGGCCGTG 759
ratTas1r2           AGCGACATCTGCATTGCCTTCCAGGAGGTTCTGCCCATACCTGAGTCCAGCCAGGTCATG 759
humanTAS1R2         CGCGACATCTGCATCGCCTTCCAGGAGACGCTGCCCACACTGCAGCCCAACCAGAACATG 747
mouseTas1r1         CGGGGCATCTGCGTCGCCTTCAAGGACGTGGTGCCTCT-------CTCCGCCCAGGCGGG 758
ratTas1r1           CGGGGCATCTGCGTCGCCTTCAAGGACATCGTGCCTTT-------CTCTGCCCGGGTGGG 752
humanTAS1R1         CAGGGGATCTGCATTGCTTTCAAGGACATCATGCCCTT-------CTCTGCCCAGGTGGG 755
mouseTas1r3         CGAGGTATCTGCATCGCACATGAGGGCCTGGTGCCACAA------CATGACACTAGTGGC 750
ratTas1r3           CGAGGTATCTGCATTGCACACGAGGGCCTGGTGCCACAA------CATGACACTAGTGGC 750
catTas1r3           AGGGGCATCTGCATCGCGCATGAGGGCCTGGTGCCACTG------C-CGCCA--GGCAGC 759
humanTAS1R3         CGCGGCATCTGCATCGCGCACGAGGGCCTGGTGCCGCTG------CCCCGTGCCGATGAC 750
                       *  ****** * **     ***       ****

mouseTas1r2         AGGCCTGAGGAGCAGGACCAACTGGACAACATCCTGGACAAGCTGCGGCGGACCTCGGCG 819
ratTas1r2           AGGTCCGAGGAGCAGAGACAACTGGACAACATCCTGGACAAGCTGCGGCGGACCTCGGCG 819
humanTAS1R2         ACGTCAGAGGAGCGCCAGCGCCTGGTGACCATTGTGGACAAGCTGCAGCAGAGCACAGCG 807
mouseTas1r1         TGACCCAAGGA-----------TGCAGCGCATGATGCTGCGTCTGGCTCGAGCCAGGACC 807
ratTas1r1           TGACCCGAGGA-----------TGCAGAGCATGATGCAGCATCTGGCTCAGGCCAGGACC 801
humanTAS1R1         CGATGAGAGGA-----------TGCAGTGCCTCATGCGCCACCTGGCCCAGGCCGGGGCC 804
mouseTas1r3         CAACAGTTGGGCAAGG------TGCTGGATGTACTACGCCAAGTGAACCAAAGTAAAGTA 804
ratTas1r3           CAACAATTGGGCAAGG------TGGTGGATGTGCTACGCCAAGTGAACCAAAGCAAAGTA 804
catTas1r3           CTGCGGCTGGGCGCCC------TACAGGGCCTGCTGCGCCAGGTGAACCAGAGCAGCGTG 813
humanTAS1R3         TCGCGGCTGGGGAAGG------TGCAGGACGTCCTGCACCAGGTGAACCAGAGCAGCGTG 804
                            **            *        *  *        **   *

mouseTas1r2         CGTGTGGTGGTGATATTCTC--GCCAGAGCTGAGCCTGCACAACTTCTTCCGCGAGGTGC 877
ratTas1r2           CGCGTCGTGGTGGTGTTCTC--GCCCGAGCTGAGCCTGTATAGCTTCTTTCACGAGGTGC 877
humanTAS1R2         CGCGTCGTGGTCGTGTTCTC--GCCCGACCTGACCCTGTACCACTTCTTCAATGAGGTGC 865
mouseTas1r1         ACCGTGGTCGTGGTCTTCTCTAACCGGCACCTGGCTGGAGTG--TTCTTCAGGTCTGTGG 865
ratTas1r1           ACCGTGGTTGTGGTCTTCTCTAACCGGCACCTGGCTAGAGTG--TTCTTCAGGTCCGTGG 859
humanTAS1R1         ACCGTCGTGGTTGTTTTTTCCAGCCGGCAGTTGGCCAGGGTG--TTTTTCGAGTCCGTGG 862
mouseTas1r3         CAAGTGGTGGTGCTGTTTGC--CTCTGCCCGTGCTGTCTACTCCCTTTTTAGTTACAGCA 862
ratTas1r3           CAGGTGGTGGTGCTGTTTGC--ATCTGCCCGTGCTGTCTACTCCCTTTTTAGCTACAGCA 862
catTas1r3           CAGGTGGTGGTGCTGTTCTC--CTCCGCCCACGCGGCCCGCACCCTCTTCAGCTACAGCA 871
humanTAS1R3         CAGGTGGTGCTGCTGTTCGC--CTCCGTGCACGCCGCCCACGCCCTCTTCAACTACAGCA 862
                       ** **  *  * **  *    * *                  * **

mouseTas1r2         TGCGCTGGAACTTCACAGGCTTTGTGTGGATTGCCTCTGAGTCCTGGGCCATCGACCCTG 937
ratTas1r2           TCCGCTGGAACTTCACGGGTTTTGTGTGGATCGCCTCTGAGTCCTGGGCTATCGACCCAG 937
humanTAS1R2         TGCGCCAGAACTTCACGGGCGCCGTGTGGATCGCCTCCGAGTCCTGGGCCATCGACCCGG 925
mouseTas1r1         TGCTGGCCAACCTGACTGGCAAAGTGTGGATCGCCTCCGAAGACTGGGCCATCTCCACGT 925
ratTas1r1           TGCTGGCCAACCTGACTGGCAAAGTGTGGGTCGCCTCAGAAGACTGGGCCATCTCCACGT 919
humanTAS1R1         TGCTGACCAACCTGACTGGCAAGGTGTGGGTCGCCTCAGAAGCCTGGGCCCTCTCCAGGC 922
mouseTas1r3         TCCATCATGGCCTCTCACCCAAGGTATGGGTGGCCAGTGAGTCTTGGCTGACATCTGACC 922
ratTas1r3           TCCTTCATGACCTCTCACCCAAGGTATGGGTGGCCAGTGAGTCCTGGCTGACCTCTGACC 922
catTas1r3           TCCGCTGCAAGCTCTCACCCAAGGTGTGGGTGGCCAGCGAGGCCTGGCTGACCTCAGACC 931
humanTAS1R3         TCAGCAGCAGGCTCTCGCCCAAGGTGTGGGTGGCCAGCGAGGCCTGGCTGACCTCTGACC 922
                    *           *  *       ** *** * ***   **    ***

mouseTas1r2         TTCTACACAACCTCACAGAGCTGCGCCACACGGGCACTTTCCTGGGCGTCACCATCCAGA 997
ratTas1r2           TTCTGCATAACCTCACGGAGCTGCGCCACACGGGTACTTTTCTGGGCGTCACCATCCAGA 997
humanTAS1R2         TCCTGCACAACCTCACGGAGCTGGGCCACTTGGGCACCTTCCTGGGCATCACCATCCAGA 985
mouseTas1r1         ACATCACCAATGTGCCCGGGATCCAGGGCATTGGGACGGTGCTGGGGGTGGCCATCCAGC 985
ratTas1r1           ACATCACCAGCGTGACTGGGATCCAAGGCATTGGGACGGTGCTCGGTGTGGCCGTCCAGC 979
humanTAS1R1         ACATCACTGGGGTGCCCGGGATCCAGCGCATTGGGATGGTGCTGGGCGTGGCCATCCAGA 982
mouseTas1r3         TGGTCATGACACTTCCCAATATTGCCCGTGTGGGCACTGTGCTTGGGTTTTTGCAGCGGG 982
ratTas1r3           TGGTCATGACACTTCCCAATATTGCCCGTGTGGGCACTGTTCTTGGGTTTCTGCAGCGCG 982
catTas1r3           TGGTCATGACGCTGCCCGGCATGCCTGGGGTGGGCACCGTGCTGGGCTTCCTGCAGCAGG 991
humanTAS1R3         TGGTCATGGGGCTGCCCGGCATGGCCCAGATGGGCACGGTGCTTGGCTTCCTCCAGAGGG 982
                       *        *  *     *          ** *   * ** **  *
```

Figure 1D

```
mouseTas1r2      GGGTGTCCATCCCTGGCTTCAGCCAGTTCCGAGTGCGCCAC---GACAAGCCAG------ 1048
ratTas1r2        GGGTGTCCATCCCTGGCTTCAGTCAGTTCCGAGTGCGCCGT---GACAAGCCAG------ 1048
humanTAS1R2      GCGTGCCCATCCCGGGCTTCAGTGAGTTCCGCGAGTGGGGC---CCACAGGCTG------ 1036
mouseTas1r1      AGAGACAAGTCCCTGGCCTGAAGGAGTTTGAAGAGTCCTAT---GTCCAGGCAGTGATGG 1042
ratTas1r1        AGAGACAAGTCCCTGGGCTGAAGGAGTTTGAGGAGTCTTAT---GTCAGGGCTGTAACAG 1036
humanTAS1R1      AGAGGGCTGTCCCTGGCCTGAAGGCGTTTGAAGAAGCCTAT---GCCCGGGCAGACAAGA 1039
mouseTas1r3      GTGCCCTACTGCCTGAATTTTCCCATTATGTGGAGACTCACCTTGCCCTGGCCGCTGACC 1042
ratTas1r3        GTGCCCTACTGCCTGAATTTTCCCATTATGTGGAGACTCGCCTTGCCCTAGCTGCTGACC 1042
catTas1r3        GCGCCCCGATGCCGGAGTTCCCATCCTACGTGCGGACCCGCCTGGCCCTGGCCGCTGACC 1051
humanTAS1R3      GTGCCCAGCTGCACGAGTTCCCCCAGTACGTGAAGACGCACCTGGCCCTGGCCACCGACC 1042
                          * *  *   *       *                        *

mouseTas1r2      -AGTATCCCATGCCTAACGAGACCAGCC---------TGAGGACTACCTGTAACCAGGAC 1098
ratTas1r2        -GGTATCCCGTGCCTAACACGACCAACC---------TGCGGACGACCTGCAACCAGGAC 1098
humanTAS1R2      -GGCCGCCACCCCTCAGCAGGACCAGCC---------AGAGCTATACCTGCAACCAGGAG 1086
mouseTas1r1      GTGCTCCCAGAACTTGCCCAGAGGGGTC---------C-TGGTGCGGCACTAACCAGCTG 1092
ratTas1r1        CTGCTCCCAGCGCTTGCCCGGAGGGGTC---------C-TGGTGCAGCACTAACCAGCTG 1086
humanTAS1R1      AGGCCCCTAGGCCTTGCCACAAGGGCTC---------C-TGGTGCAGCAGCAATCAGCTC 1089
mouseTas1r3      CAGCATTCTGTGCCTCACTGAATGCGGA---GTTGGATCTGGAGGAACATGTGATGGGGC 1099
ratTas1r3        CAACATTCTGTGCCTCCCTGAAAGCTGA---GTTGGATCTGGAGGAGCGCGTGATGGGGC 1099
catTas1r3        CTGCCTTCTGCGCCTCGCTGGACGCTGAACAGCCAGGCCTGGAGGAGCACGTGGTGGGGC 1111
humanTAS1R3      CGGCCTTCTGCTCTGCCCTGGGCGAGAGGGAGCAGGGTCTGGAGGAGGACGTGGTGGGCC 1102
                             *    *                      *               *

mouseTas1r2      TGTGACGCCTGCATGAACATCACCGAGTCCTTTAACAACGTTC----------------- 1141
ratTas1r2        TGTGACGCCTGCTTGAACACCACCAAGTCCTTCAACAACATCC----------------- 1141
humanTAS1R2      TGCGACAACTGCCTGAACGCCACCTTGTCCTTCAACACCATTC----------------- 1129
mouseTas1r1      TGCAGGGAGTGTCACGCTTTCACGACATGGAACATGCCCGAGC----------------- 1135
ratTas1r1        TGCCGGGAGTGCCACACGTTCACGACTCGTAACATGCCCACGC----------------- 1129
humanTAS1R1      TGCAGAGAATGCCAAGCTTTCATGGCACACACGATGCCCAAGC----------------- 1132
mouseTas1r3      AACGCTGTCCACGGTGTGACGACATCATGCTGCAGAACCTATCATCTGGGCTGTTGCAGA 1159
ratTas1r3        CACGCTGTTCACAATGTGACTACATCATGCTACAGAACCTGTCATCTGGGCTGATGCAGA 1159
catTas1r3        CACGCTGCCCCCAATGTGACCACGTCACGCTAGAGAACCT-------------------- 1151
humanTAS1R3      AGCGCTGCCCGCAGTGTGACTGCATCACGCTGCAGAACGT-------------------- 1142
                                                  *

mouseTas1r2      -TCATGCTTTCGGGGGAGCGTGTGGTC-----TACAGTGTGTACTCGGCCGTCTACGCGG 1195
ratTas1r2        -TTATACTTTCGGGGGAGCGCGTGGTC-----TACAGCGTGTACTCGGCAGTTTACGCGG 1195
humanTAS1R2      -TCAGGCTCTCTGGGGAGCGTGTCGTC-----TACAGCGTGTACTCTGCGGTCTATGCTG 1183
mouseTas1r1      -TTGGAGCCTTCTCCATGAGCGCTGCC-----TACAATGTGTATGAGGCTGTGTATGCTG 1189
ratTas1r1        -TTGGAGCCTTCTCCATGAGTGCCGCC-----TACAGAGTGTATGAGGCTGTGTACGCTG 1183
humanTAS1R1      -TCAAAGCCTTCTCCATGAGTTCTGCC-----TACAACGCATACCGGGCTGTGTATGCGG 1186
mouseTas1r3      ACCTATCAGCTGGGCAATTGCACCACCAAATATTTGCAACCTATGCAGCTGTGTACAGTG 1219
ratTas1r3        ACCTATCAGCTGGGCAGTTGCACCACCAAATATTTGCAACCTATGCAGCTGTGTACAGTG 1219
catTas1r3        ----ATCTGCGGGGCTGCTGCACCACCAGACCTTCGCTGCCTACGCGGCTGTGTATGGCG 1207
humanTAS1R3      ----GAGCGCAGGGCTAAATCACCACCAGACGTTCTCTGTCTACGCAGCTGTGTATAGCG 1198
                                           *     *        **    ** ** **    *

mouseTas1r2      TAGCCCACACCCTCCACAGACTCCTCCACTGCAACCAGGTCCGCTGCACCA---AGCAAA 1252
ratTas1r2        TGGCCCATGCCCTCCACAGACTCCTCGGCTGTAACCGGGTCCGCTGCACCA---AGCAAA 1252
humanTAS1R2      TGGCCCATGCCCTGCACAGCCTCCTCGGCTGTGACAAAAGCACCTGCACCA---AGAGGG 1240
mouseTas1r1      TGGCCCACGGCCTCCACCAGCTCCTGGGATGTACCTCTGGACCTGTGCCA---GAGGCC 1246
ratTas1r1        TGGCCCACGGCCTCCACCAGCTCCTGGGATGTACTTCTGAGATCTGTTCCA---GAGGCC 1240
humanTAS1R1      TGGCCCATGGCCTCCACCAGCTCCTGGGCTGTGCCTCTGGAGCTTGTTCCA---GGGGCC 1243
mouseTas1r3      TGGCTCAAGCCCTTCACAACACCCTACAGTGCAATGTCTCACATTGCCACGTATCAGAAC 1279
ratTas1r3        TGGCTCAGGCCCTTCACAACACCCTGCAGTGCAATGTCTCACATTGCCACACATCAGAGC 1279
catTas1r3        TGGCCCAAGCCCTTCACAACACACTGCGCTGCAATGCCTCGGGCTGCCCCAGGCGGGAGC 1267
humanTAS1R3      TGGCCCAGGCCCTGCACAACACTCTTCAGTGCAACGCCTCAGGCTGCCCCGCGCAGGACC 1258
                 * ** **   *** ***      **    **              **  *

mouseTas1r2      TCGTCTATCCATGGCAGCTACTCAGGGAGATCTGGCATGTCAACTTCACGCTCCTGGGCA 1312
ratTas1r2        AGGTCTACCCGTGGCAGCTACTCAGGGAGATCTGGCACGTCAACTTCACGCTCCTGGGTA 1312
humanTAS1R2      TGGTCTACCCCTGGCAGCTGCTTGAGGAGATCTGGAAGGTCAACTTCACTCTCCTGGACC 1300
mouseTas1r1      CAGTCTACCCCTGGCAGCTTCTTCAGCAGATCTACAAGGTGAATTTCCTTCTACATAAGA 1306
ratTas1r1        CAGTCTACCCCTGGCAGCTTCTTCAGCAGATCTACAAGGTGAATTTTCTTCTACATGAGA 1300
humanTAS1R1      GAGTCTACCCCTGGCAGCTTTTGGAGCAGATCCACAAGGTGCATTTCCTTCTACACAAGG 1303
mouseTas1r3      ATGTTCTACCCTGGCAGCTCCTGGAGAACATGTACAATATGAGTTTCCATGCTCGAGACT 1339
ratTas1r3        CTGTTCAACCCTGGCAGCTCCTGGAGAACATGTACAATATGAGTTTCCGTGCTCGAGACT 1339
catTas1r3        CTGTGCGGCCCTGGCAGCTCCTAGAGAACATGTACAACGTGAGCTTCCGTGCTCGCGGCC 1327
humanTAS1R3      CCGTGAAGCCCTGGCAGCTCCTGGAGAACATGTACAACCTGACCTTCCACGTGGGCGGGC 1318
                   **    ** ********  *   * * **     *  *    **
```

Figure 1E

```
mouseTas1r2        ACCAGCTCTTCTTCGACGAACAAGGGGACATGCCGATGCTCCTGGACATCATCCAGTGGC 1372
ratTas1r2          ACCGGCTCTTCTTTGACCAACAAGGGGACATGCCGATGCTCTTGGACATCATCCAGTGGC 1372
humanTAS1R2        ACCAAATCTTCTTCGACCCGCAAGGGGACGTGGCTCTGCACTTGGAGATTGTCCAGTGGC 1360
mouseTas1r1        AGACTGTAGCATTCGATGACAAGGGGGACCCTCTAGGTTATTATGACATCATCGCCTGGG 1366
ratTas1r1          ATACTGTGGCATTTGATGACAACGGGGACACTCTAGGTTACTACGACATCATCGCCTGGG 1360
humanTAS1R1        ACACTGTGGCGTTTAATGACAACAGAGATCCCCTCAGTAGCTATAACATAATTGCCTGGG 1363
mouseTas1r3        TGACACTACAGTTTGATGCTGAAGGGAATGTAGACATGGAATATGACCTGAAGATGTGGG 1399
ratTas1r3          TGACACTGCAGTTTGATGCCAAAGGGAGTGTAGACATGGAATATGACCTGAAGATGTGGG 1399
catTas1r3          TGGCACTGCAGTTCGACGCCAGCGGGAACGTGAACGTGGATTACGACCTGAAACTGTGGG 1387
humanTAS1R3        TGCCGCTGCGGTTCGACAGCAGCGGAAACGTGGACATGGAGTACGACCTGAAGCTGTGGG 1378
                         *    **  *        *                    *  *       ***

mouseTas1r2        AATGGGGCCTGAGCCAGAACCCCTTCCAAAGCATCGCCTCCTACTCCCCCACCGAGACGA 1432
ratTas1r2          AGTGGGACCTGAGCCAGAATCCCTTCCAAAGCATCGCCTCCTATTCTCCCACCAGCAAGA 1432
humanTAS1R2        AATGGGACCGGAGCCAGAATCCCTTCCAGAGCGTCGCCTCCTACTACCCCCTGCAGCGAC 1420
mouseTas1r1        ACTGGAATGGACCTGAATGGACCTTTGAGGTCATTGGTTCTGCCTCACTGTCTCCAGTTC 1426
ratTas1r1          ACTGGAATGGACCTGAATGGACCTTTGAGATCATTGGCTCTGCCTCACTGTCTCCAGTTC 1420
humanTAS1R1        ACTGGAATGGACCCAAGTGGACCTTCACGGTCCTCGGTTCCTCCACATGGTCTCCAGTTC 1423
mouseTas1r3        TGTGGCAGAGCCCTACACCTGTATTACATACTGTGGCACCT---------TCAACGGCA 1450
ratTas1r3          TGTGGCAGAGCCCTACACCTGTACTACATACTGTAGGCACCT---------TCAACGGCA 1450
catTas1r3          TGTGGCAGGACCCGACGCCCGAGCTGCGCACCGTAGGCACCT---------TCAAGGGCC 1438
humanTAS1R3        TGTGGCAGGGCTCAGTGCCCAGGCTCCACGACGTGGGCAGGT---------TCAACGGCA 1429
                     ***                   *        * *

mouseTas1r2        GGCT-GACCTACATTAGCAATGTGTCC--TGGTACACCCCCAACAACACGGTCCCCATAT 1489
ratTas1r2          GGCT-AACCTACATTAACAATGTGTCC--TGGTACACCCCCAACAACACGGTCCCTGTCT 1489
humanTAS1R2        AGCT-GAAGAACATCCAAGACATCTCC--TGGCACACCGTCAACAACACGATCCCTATGT 1477
mouseTas1r1        ATCTAGACATAAATAAGACAAAAATCCAGTGGCACGGGAAGAACAATCAGGTGCCTGTGT 1486
ratTas1r1          ATCTGGACATAAATAAGACAAAAATCCAGTGGCACGGGAAGAACAATCAGGTGCCTGTGT 1480
humanTAS1R1        AGCTAAACATAAATGAGACCAAAATCCAGTGGCACGGAAAGGACAACCAGGTGCCTAAGT 1483
mouseTas1r3        CCCTTCAGCTGCAGCAGTCTAAAATGTACTGGC------CAGGCAACCAGGTGCCAGTCT 1504
ratTas1r3          CCCTTCAGCTGCAGCACTCGAAAATGTATTGGC------CAGGCAACCAGGTGCCAGTCT 1504
catTas1r3          GCCTGGAGCTCTGGCGCTCTCAGATGTGCTGGCACACGCCGGGGAAGCAGCAGCCCGTGT 1498
humanTAS1R3        GCCTCAGGACAGAGCGCCTGAAGATCCGCTGGCACACGTCTGACAACCAGAAGCCCGTGT 1489
                     **                    *    ***            **   *   **    *

mouseTas1r2        CCATGTGTTCTAAGAGTTGCCAGCCTGGGCAAATGAAAAAACCCATAGGCCTCCACCCGT 1549
ratTas1r2          CCATGTGTTCCAAGAGCTGCCAGCCAGGGCAAATGAAAAAGTCTGTGGGCCTCCACCCTT 1549
humanTAS1R2        CCATGTGTTCCAAGAGGTGCCAGTCAGGGCAAAAGAAGAAGCCTGTGGGCATCCACGTCT 1537
mouseTas1r1        CAGTGTGTACCAGGGACTGTCTCGAAGGGCACCACAGGTTGGTCATGGGTTCCCACCACT 1546
ratTas1r1          CAGTGTGTACCACGGACTGTCTGGCAGGGCACCACAGGGTGGTTGTGGGTTCCCACCACT 1540
humanTAS1R1        CTGTGTGTTCCAGCGACTGTCTTGAAGGGCACCAGCGAGTGGTTACGGGTTTCCATCACT 1543
mouseTas1r3        CCCAGTGTTCCCGCCAGTGCAAAGATGGCCAGGTTCGCCGAGTAAAGGGCTTTCATTCCT 1564
ratTas1r3          CCCAGTGCTCCCGGCAGTGCAAAGATGGCCAGGTGCGCAGAGTAAAGGGCTTTCATTCCT 1564
catTas1r3          CCCAGTGCTCCCGGCAGTGCAAGGAAGGCCAGGTGCGCCGCGTGAAGGGCTTCCACTCTT 1558
humanTAS1R3        CCCGGTGCTCGCGGCAGTGCCAGGAGGGCCAGGTGCGCCGGGTCAAGGGGTTCCACTCCT 1549
                   *   ***  *       **       ** **                **    **    *

mouseTas1r2        GCTGCTTCGAGTGTGTGGACTGTCCGCCGGGCACCTACCTCAACCGATCAGTAGATGAGT 1609
ratTas1r2          GTTGCTTCGAGTGCTTGGATTGTATGCCAGGCACCTACCTCAACCGCTCAGCAGATGAGT 1609
humanTAS1R2        GCTGCTTCGAGTGCATCGACTGCCTTCCCGGCACCTTCCTCAACCACACTGAAGATGAAT 1597
mouseTas1r1        GCTGCTTCGAGTGCATGCCCTGTGAAGCTGGGACATTTCTCAAC---ACGAGTGAGCTTC 1603
ratTas1r1          GCTGCTTTGAGTGTGTGCCCTGCGAAGCTGGGACCTTTCTCAAC---ATGAGTGAGCTTC 1597
humanTAS1R1        GCTGCTTTGAGTGTGTGCCCTGTGGGGCTGGGACCTTCCTCAAC---AAGAGTGACCTCT 1600
mouseTas1r3        GCTGCTATGACTGCGTGGACTGCAAGGCGGGCAGCTACCGGAAG---CATCCAGATGACT 1621
ratTas1r3          GCTGCTATGACTGTGTGGACTGCAAGGCAGGGAGCTACCGGAAG---CATCCAGATGACT 1621
catTas1r3          GCTGTTACAACTGCGTGGACTGCAAGGCGGGCAGTTATCAGCGC---AACCCAGATGACC 1615
humanTAS1R3        GCTGCTACGACTGTGTGGACTGCGAGGCGGGCAGCTACCGGCAA---AACCCAGACGACA 1606
                   * ** *   * **  *    **     * ** *  *  *              **

mouseTas1r2        TTAACTGTCTGTCCTGCCCGGGTTCCATGTGGTCTTACAAGAACAACATCGCTTGCTTCA 1669
ratTas1r2          TTAACTGTCTGTCCTGCCCGGGTTCCATGTGGTCCTACAAGAACGACATCACTTGCTTCC 1669
humanTAS1R2        ATGAATGCCAGGCCTGCCCGAATAACGAGTGGTCCTACCAGAGTGAGACCTCCTGCTTCA 1657
mouseTas1r1        ACACCTGCCAGCCTTGTGGAACAGAAGAATGGGCCCCTGAGGGGAGCTCAGCCTGCTTCT 1663
ratTas1r1          ACATCTGCCAGCCTTGTGGAACAGAAGAATGGGCACCCAAGGAGAGCACTACTTGCTTCC 1657
humanTAS1R1        ACAGATGCCAGCCTTGTGGGAAAGAAGAGTGGGCACCTGAGGGAAGCCAGACCTGCTTCC 1660
mouseTas1r3        TCACCTGTACTCCATGTAACCAGGACCAGTGGTCCCCAGAGAAAAGCACAGCCTGCTTAC 1681
ratTas1r3          TCACCTGTACTCCATGTGGCAAGGATCAGTGGTCCCCAGAAAAAAGCACAACCTGCTTAC 1681
catTas1r3          TCCTCTGCACCCAGTGTGACCAGGACCAGTGGTCCCCAGACCGGAGCACACGCTGCTTCG 1675
humanTAS1R3        TCGCCTGCACCTTTTGTGGCCAGGATGAGTGGTCCCCGGAGCGAAGCACACGCTGCTTCC 1666
                        **       **             *** *     *             *****
```

Figure 1F

```
mouseTas1r2         AGCGGCGGCTGGCCTTCCTGGAGTGGCACGAAGTGCCCACTATCGTGGTGACCATCCTGG 1729
ratTas1r2           AGCGGCGGCCTACCTTCCTGGAGTGGCACGAAGTGCCCACCATCGTGGTGGCCATACTGG 1729
humanTAS1R2         AGCGGCAGCTGGTCTTCCTGGAATGGCATGAGGCACCCACCATCGCTGTGGCCCTGCTGG 1717
mouseTas1r1         CACGCACCGTGGAGTTCTTGGGGTGGCATGAACCCATCTCTTTGGTGCTATTAGCAGCTA 1723
ratTas1r1           CACGCACGGTGGAGTTCTTGGCTTGGCATGAACCCATCTCTTTGGTGCTAATAGCAGCTA 1717
humanTAS1R1         CGCGCACTGTGGTGTTTTTGGCTTTGCGTGAGCACACCTCTTGGGTGCTGCTGGCAGCTA 1720
mouseTas1r3         CTCGCAGGCCCAAGTTTCTGGCTTGGGGGAGCCAGTTGTGCTGTCACTCCTCCTGCTGC 1741
ratTas1r3           CTCGCAGGCCCAAGTTTCTGGCTTGGGGGAGCCAGCTGTGCTGTCACTTCTCCTGCTGC 1741
catTas1r3           CCCGCAAGCCCATGTTCCTGGCATGGGGGAGCCAGCTGTGCTGCTACTGCTCGCGCTGC 1735
humanTAS1R3         GCCGCAGGTCTCGGTTCCTGGCATGGGGCGAGCCGGCTGTGCTGCTGCTGCTCCTGCTGC 1726
                      **          **  ***  * *   **                 *

mouseTas1r2         CCGCCCTGG-GCTTCATCAGTACGCTGGCCATTCTGCTCATCTTCTGGAGACATTTCCAG 1788
ratTas1r2           CTGCCCTGG-GCTTCTTCAGTACACTGGCCATTCTTTTCATCTTCTGGAGACATTTCCAG 1788
humanTAS1R2         CCGCCCTGG-GCTTCCTCAGCACCCTGGCCATCCTGGTGATATTCTGGAGGCACTTCCAG 1776
mouseTas1r1         ACACGCTATTGCTGCTGCTGCTGATTGGGACTGCTGGCC-TGTTTGCCTGGCGTCTTCAC 1782
ratTas1r1           ACACGCTATTGCTGCTGCTGCTGGTTGGGACTGCTGGCC-TGTTTGCCTGGCATTTTCAC 1776
humanTAS1R1         ACACGCTGCTGCTGCTGCTGCTGCTTGGGACTGCTGGCC-TGTTTGCCTGGCACCTAGAC 1779
mouseTas1r3         TTTGCCTGGTGCTGGGTCTA-GCACTGGCTGCTCTGGGGCTCTCTGTCCACCACTGGGAC 1800
ratTas1r3           TTTGCCTGGTGCTGGGCCTG-ACACTGGCTGCCCTGGGGCTCTTTGTCCACTACTGGGAC 1800
catTas1r3           TGGCTCTGGCGCTGGGCCTG-GCGCTGGCAGCCCTGGGGCTCTTCCTCTGGCACTCGGAC 1794
humanTAS1R3         TGAGCCTGGCGCTGGGCCTT-GTGCTGGCTGCTTTGGGGCTGTTCGTTCACCATCGGGAC 1785
                         **   ***    *       ***      *     * *.              *

mouseTas1r2         ACGCCCATGGTGCGCTCGGCGGGCGGCCCCATGTGCTTCCTGATGCTGGTGCCCCTGCTG 1848
ratTas1r2           ACACCCATGGTGCGCTCGGCCGGTGGCCCCATGTGCTTCCTGATGCTCGTGCCCCTGCTG 1848
humanTAS1R2         ACACCCATAGTTCGCTCGGCTGGGGGCCCCATGTGCTTCCTGATGCTGACACTGCTGCTG 1836
mouseTas1r1         ACGCCTGTTGTGAGGTCAGCTGGGGGTAGGCTGTGCTTCCTCATGCTGGGTTCCTTGGTA 1842
ratTas1r1           ACACCTGTAGTGAGGTCAGCTGGGGGTAGGCTGTGCTTCCTCATGCTGGGTTCCCTGGTG 1836
humanTAS1R1         ACCCCTGTGGTGAGGTCAGCAGGGGGCCGCCTGTGCTTTCTTATGCTGGGCTCCCTGGCA 1839
mouseTas1r3         AGCCCTCTTGTCCAGGCCTCAGGTGGCTCACAGTTCTGCTTTGGCCTGATCTGCCTAGGC 1860
ratTas1r3           AGCCCTCTTGTTCAGGCCTCAGGTGGGTCACTGTTCTGCTTTGGCCTGATCTGCCTAGGC 1860
catTas1r3           AGCCCGCTGGTTCAGGCCTCAGGTGGCCACGGGCCTGCTTTGGCCTGGCTTGCCTGGGC 1854
humanTAS1R3         AGCCCACTGGTTCAGGCCTCGGGGGGCCCCTGGCCTGCTTTGGCCTGGTGTGCCTGGGC 1845
                    *  **  * **     *  * ** **      *  **   *    **        *

mouseTas1r2         CTGGCGTTCGGGATGGTCCCCGTGTATGTGGCCCCCCCACGGTCTTCTCCTGTTTCTGC 1908
ratTas1r2           CTGGCGTTTGGGATGGTGCCCGTGTATGTGGGGCCCCCCACGGTCTTCTCATGCTTCTGC 1908
humanTAS1R2         GTGGCATACATGGTGGTCCCGGTGTACGTGGGGCCGCCCAAGGTCTCCACCTGCCTCTGC 1896
mouseTas1r1         GCTGGGAGTTGCAGCCTCTACAGCTTCTTCGGGAAGCCCACGGTGCCCGCGTGCTTGCTG 1902
ratTas1r1           GCCGGAAGTTGCAGCTTCTATAGCTTCTTCGGGGAGCCCACGGTGCCCGCGTGCTTGCTG 1896
humanTAS1R1         GCAGGTAGTGGCAGCCTCTATGGCTTCTTTGGGGAACCCACAAGGCCTGCGTGCTTGCTA 1899
mouseTas1r3         CTCTTCTGCCTCAGTGTCCTTCTGTTCCCAGGGCGGCCAAGCTCTGCCAGCTGCCTTGCA 1920
ratTas1r3           CTCTTCTGCCTCAGTGTCCTTCTGTTCCCAGGACGACCACGCTCTGCCAGCTGCCTTGCC 1920
catTas1r3           CTGGTCTGCCTCAGTGTCCTCCTGTTCCCTGGCCAGCCAGGCCCTGCCAGCTGCCTGGCC 1914
humanTAS1R3         CTGGTCTGCCTCAGCGTCCTCCTGTTCCCTGGCCAGCCCAGCCCTGCCCGATGCCTGGCC 1905
                                    *       *     **    **             **  *

mouseTas1r2         CGCCAGGCTTTCTTCACCGTTTGCTTCTCCGTCTGCCTCTCCTGCATCACGGTGCGCTCC 1968
ratTas1r2           CGACAGGCTTTCTTCACCGTCTGCTTCTCCATCTGCCTATCCTGCATCACCGTGCGCTCC 1968
humanTAS1R2         CGCCAGGCCCTCTTTCCCCTCTGCTTCACAATTTGCATCTCCTGTATCGCCGTGCGTTCT 1956
mouseTas1r1         CGTCAGCCCCTCTTTTCTCTCGGGTTTGCCATTTTCCTCTCCTGTCTGACAATCCGCTCC 1962
ratTas1r1           CGTCAGCCCCTCTTTTCTCTCGGGTTTGCCATCTTCCTCTCCTGCCTGACAATCCGCTCC 1956
humanTAS1R1         CGCCAGGCCCTCTTTGCCCTTGGTTTCACCATCTTCCTGTCCTGCCTGACAGTTCGCTCA 1959
mouseTas1r3         CAACAACCAATGGCTCACCTCCCTCTCACAGGCTGCCTGAGCACACTCTTCCTGCAAGCA 1980
ratTas1r3           CAACAACCAATGGCTCACCTCCCTCTCACAGGCTGCCTGAGCACACTCTTCCTGCAAGCA 1980
catTas1r3           CAGCAGCCACTGTTCCACCTCCCACTCACTGGCTGCCTGAGCACGTTTTTCCTGCAAGCG 1974
humanTAS1R3         CAGCAGCCCTTGTCCCACCTCCCGCTCACGGGCTGCCTGAGCACACTCTTCCTGCAGGCG 1965
                    *  **  *  *        *     *  *    * * *   *    *     * *   *

mouseTas1r2         TTCCAGATTGTGTGCGTCTTCAAGATGGCCAGACGCCTGCCAAGCGCCTACGGTTTCTGG 2028
ratTas1r2           TTCCAGATCGTGTGTGTCTTCAAGATGGCCAGACGCCTGCCAAGTGCCTACAGTTTTTGG 2028
humanTAS1R2         TTCCAGATCGTCTGCGCCTTCAAGATGGCCAGCCGCTTCCCACGCGCCTACAGCTACTGG 2016
mouseTas1r1         TTCCAACTGGTCATCATCTTCAAGTTTTCTACCAAGGTACCCACATTCTACCACACTTGG 2022
ratTas1r1           TTCCAACTGGTCATCATCTTCAAGTTTTCTACCAAGGTGCCCACATTCTACCGTACCTGG 2016
humanTAS1R1         TTCCAACTAATCATCATCTTCAAGTTTTCCACCAAGGTACCTACATTCTACCACGCCTGG 2019
mouseTas1r3         GCTGAGACCTTTGTGGAGTCTGAGCTGCCACTGAGCTGGGCAAACTGGCTATGCAGCTAC 2040
ratTas1r3           GCCGAGATCTTTGTGGAGTCTGAGCTGCCACTGAGTTGGGCAAACTGGCTCTGCAGCTAC 2040
catTas1r3           GCCGAGATATTTGTGGGGTCGGAGCTGCCACCAAGCTGGGCTGAGAAGATGCGTGGCCGC 2034
humanTAS1R3         GCCGAGATCTTCGTGGAGTCAGAACTGCCTCTGAGCTGGGCAGACCGGCTGAGTGGCTGC 2025
                        *     *       *   *  *  *           *
```

Figure 1G

```
mouseTas1r2    ATGCGTTACCACGGGCCCTACGTCTTTGTGGCCTTCATCACGGCCGTCAAGGTGGCCCT- 2087
ratTas1r2      ATGCGTTACCACGGGCCCTATGTCTTCGTGGCCTTCATCACGGCCATCAAGGTGGCCCT- 2087
humanTAS1R2    GTCCGCTACCAGGGGCCCTACGTCTCTATGGCATTTATCACGGTACTCAAAATGGTCAT- 2075
mouseTas1r1    GCCCAAAACCATGGTGCCGGAATATTCGTCATTGTCAGCTCCACGGTCCATTTGTTCCTC 2082
ratTas1r1      GCCCAAAACCATGGTGCAGGTCTATTCGTCATTGTCAGCTCCACGGTCCATTTGCTCATC 2076
humanTAS1R1    GTCCAAAACCACGGTGCTGGCCTGTTTGTGATGATCAGCTCAGCGGCCCAGCTGCTTATC 2079
mouseTas1r3    CTTCGGGGACTCTGGGCCTGGCTAGTGGTACTGTTGGCCACTTTTGTGGAGGCAGCACTA 2100
ratTas1r3      CTTCGGGGCCCCTGGGCTTGGCTGGTGGTACTGCTGGCCACTCTTGTGGAGGCTGCACTA 2100
catTas1r3      CTGCGGGGGCCCTGGGCCTGGCTGGTGGTGCTGCTTGCTATGCTGGCAGAAGCCGCATTG 2094
humanTAS1R3    CTGCGGGGGCCCTGGGCCTGGCTGGTGGTGCTGCTGGCCATGCTGGTGGAGGTCGCACTG 2085
                  *     *   *  *     *     *     *              *        *

mouseTas1r2    GGTGGCAGGCAACATGCTGGCCACCACCATCAACCCCATTGGCCGGACCGACCCCGATGA 2147
ratTas1r2      GGTGGTGGGCAACATGCTGGCCACCACCATCAACCCCATTGGCCGGACCGACCCGGATGA 2147
humanTAS1R2    TGTGGTAATTGGCATGCTGGCCACGGGCCTCAGTCCCACCACCCGTACTGACCCCGATGA 2135
mouseTas1r1    TGTCTCACGTGGCTTGCAATGTGGACCCCACGGCCCACCAGGGAGTACCAGCGCTT---- 2138
ratTas1r1      TGTCTCACATGGCTTGTAATGTGGACCCCACGACCCACCAGGGAATACCAGCGCTT---- 2132
humanTAS1R1    TGTCTAACTTGGCTGGTGGTGTGGACCCCACTGCCTGCTAGGGAATACCAGCGCTT---- 2135
mouseTas1r3    TGTGCCTGGTATTTGATCGCTTTCCCACCAGAGGTGGT---GACAGACTGGTCAGTGCTG 2157
ratTas1r3      TGTGCCTGGTACTTGATGGCTTTCCCTCCAGAGGTGGT---GACAGATTGGCAGGTGCTG 2157
catTas1r3      TGTGCCTGGTACCTGGTAGCCTTCCCGCCAGAGGTGGT---GACGGACTGGCGGGTACTG 2151
humanTAS1R3    TGCACCTGGTACCTGGTGGCCTTCCCGCCGGAGGTGGT---GACGGACTGGCACATGCTG 2142
                *                         *                  *

mouseTas1r2    CCCCAATATCATAATCCTCTCCTGCCACCCTAACTACCGCAACGGGCTACTCTTCAACAC 2207
ratTas1r2      CCCCAACATCATGATCCTCTCGTGCCACCCTAACTACCGCAACGGGCTACTGTTCAACAC 2207
humanTAS1R2    CCCCAAGATCACAATTGTCTCCTGTAACCCCAACTACCGCAACAGCCTGCTGTTCAACAC 2195
mouseTas1r1    CCCCCATCTGGTGATTCTTGAGTGC-ACAGAGGTCAACTCTGTGGGCTTCCTGGTGGCTT 2197
ratTas1r1      CCCCCATCTGGTGATTCTCGAGTGC-ACAGAGGTCAACTCTGTAGGCTTCCTGTTGGCTT 2191
humanTAS1R1    CCCCCATCTGGTGATGCTTGAGTGC-ACAGAGACCAACTCCCTGGGCTTCATACTGGCCT 2194
mouseTas1r3    CCCACAGA-GGTACTGGAGCACTGCCACGTGCGTT-CCTGGGTCAGCCTGGGCTTGGTGC 2215
ratTas1r3      CCCACGGA-GGTACTGGAACACTGCCGCATGCGTT-CCTGGGTCAGCCTGGGCTTGGTGC 2215
catTas1r3      CCCACAGA-GGCGCTGGTGCACTGCCACGTGCACT-CCTGGATCAGCTTCGGCCTGGTGC 2209
humanTAS1R3    CCCACGGA-GGCGCTGGTGCACTGCCGCACACGCT-CCTGGGTCAGCTTCGGCCTAGCGC 2200
               ***           *       **   *         *        *

mouseTas1r2    CAGCATG-GACTTGCTGCTGTCCGTGCTGGGTTTCAGCTTCGCGTACGTGGGCAAGGAAC 2266
ratTas1r2      CAGCATG-GACTTGCTGCTGTCTGTGCTGGGTTTCAGCTTCGCTTACATGGGCAAGGAGC 2266
humanTAS1R2    CAGCCTG-GACCTGCTGCTCTCAGTGGTGGGTTTCAGCTTCGCCTACATGGGCAAAGAGC 2254
mouseTas1r1    TCGCACACAACATCCTCCTCTCCATCAGCACCTTTGTCTGCAGCTACCTGGGTAAGGAAC 2257
ratTas1r1      TCACCCACAACATTCTCCTCTCCATCAGTACCTTCGTCTGCAGCTACCTGGGTAAGGAAC 2251
humanTAS1R1    TCCTCTACAATGGCCTCCTCTCCATCAGTGCCTTTGCCTGCAGCTACCTGGGTAAGGACT 2254
mouseTas1r3    ACATCACCAATGCAATGTTAGCTTTCCTCTGCTTTCTGGGCACTTTCCTGGTACAGAGCC 2275
ratTas1r3      ACATCACCAATGCAGTGTTAGCTTTCCTCTGCTTTCTGGGCACTTTCCTGGTACAGAGCC 2275
catTas1r3      ATGCCACTAACGCCATGCTGGCCTTCCTCTGCTTCCTGGGCACTTTCCTGGTGCAGAGCC 2269
humanTAS1R3    ACGCCACCAATGCCACGCTGGCCTTTCTCTGCTTCCTGGGCACTTTCCTGGTGCGGAGCC 2260
                        *        *  *  *       **      *   * * ***

mouseTas1r2    TGCCCACCAACTACAACGAAGCCAAGTTCATCACCCTCAGCATGACCTTCTCCTTCACCT 2326
ratTas1r2      TGCCCACCAACTACAACGAAGCCAAGTTCATCACTCTCAGCATGACCTTCTCCTTCACCT 2326
humanTAS1R2    TGCCCACCAACTACAACGAGGCCAAGTTCATCACCCTCAGCATGACCTTCTATTTCACCT 2314
mouseTas1r1    TGCCGGAGAACTATAACGAAGCCAAATGTGTCACCTTCAGCCTGCTCCTCCACTTCGTAT 2317
ratTas1r1      TGCCAGAGAACTATAATGAAGCCAAATGTGTCACCTTCAGCCTGCTCCTCAACTTCGTAT 2311
humanTAS1R1    TGCCAGAGAACTACAACGAGGCCAAATGTGTCACCTTCAGCCTGCTCTTCAACTTCGTGT 2314
mouseTas1r3    AGCCTGGCCGCTACAACCGTGCCCGTGGTCTCACCTTCGCCATGCTAGCTTATTTCATCA 2335
ratTas1r3      AGCCTGGTCGCTATAACCGTGCCCGTGGCCTCACCTTCGCCATGCTAGCTTATTTCATCA 2335
catTas1r3      GGCCAGGCCGCTACAATGGTGCCCGCGGCCTCACCTTTGCCATGCTGGCCTACTTCATCA 2329
humanTAS1R3    AGCCGGGCCGCTACAACCGTGCCCGTGGCCTCACCTTTGCCATGCTGGCCTACTTCATCA 2320
                ***      *** **    ***       ****  *   * **         ***

mouseTas1r2    CCTCCATCTCCCTCTGCACGTTCATGTCTGTCCACGATGGCGTGCTGGTCACCATCATGG 2386
ratTas1r2      CCTCCATCTCCCTCTGCACCTTCATGTCTGTGCACGACGGCGTGCTGGTCACCATCATGG 2386
humanTAS1R2    CATCCGTCTCCCTCTGCACCTTCATGTCTGCCTACAGCGGGGTGCTGGTCACCATCGTGG 2374
mouseTas1r1    CCTGGATCGCTTTCTTCACCATGTCCAGCATTTACCAGGGCA-GCTACCTACC--CGCGG 2374
ratTas1r1      CCTGGATCGCCTTCTTCACCATGGCCAGCATTTACCAGGGCA-GCTACCTGCC--TGCGG 2368
humanTAS1R1    CCTGGATCGCCTTCTTCACCACGGCCAGCGTCTACGACGGCA-AGTACCTGCC--TGCGG 2371
mouseTas1r3    CCTGGGTCTCTTTTGTGCCCCTCCTGGCCA-ATGTGCAGGTGGCCTACCAGCC--AGCTG 2392
ratTas1r3      TCTGGGTCTCTTTTGTGCCCCTCCTGGCTA-ATGTGCAGGTGGCCTACCAGCC--AGCTG 2392
catTas1r3      CCTGGATCTCCTTTGTGCCCCTCTTTGCCA-ATGTGCACGTGGCCTACCAGCC--TGCCG 2386
humanTAS1R3    CCTGGGTCTCCTTTGTGCCCCTCCTGGCCA-ATGTGCAGGTGGTCCTCAGGCC--CGCCG 2377
                 *   ** *  *     *                    *           **      *
```

Figure 1H

```
mouseTas1r2      ATCTCCTGGTCACTGTGCTCAACTTTCTGGCCATCGGCTTG---GGGTACTTTGGCCCCA 2443
ratTas1r2        ACCTCCTGGTCACTGTGCTCAACTTCCTGGCCATCGGCTTG---GGATACTTTGGCCCCA 2443
humanTAS1R2      ACCTCTTGGTCACTGTGCTCAACCTCCTGGCCATCAGCCTG---GGCTACTTCGGCCCCA 2431
mouseTas1r1      TCAATGTGCTGGCAGGGCTGGCCACTCTGAGTGGCGGCTTCAGCGGCTATTTCCTCCCTA 2434
ratTas1r1        TCAATGTGCTGGCAGGGCTGACCACACTGAGCGGCGGCTTCAGCGGTTACTTCCTCCCCA 2428
humanTAS1R1      CCAACATGATGGCTGGGCTGAGCAGCCTGAGCAGCGGCTTCGGTGGGTATTTTCTGCCTA 2431
mouseTas1r3      TGCAGATGGGTGCTATCCTAGTCTGTGCCCTGGGCATCCTGGTCACCTTCCACCTGCCCA 2452
ratTas1r3        TGCAGATGGGTGCTATCTTATTCTGTGCCCTGGGCATCCTGGCCACCTTCCACCTGCCCA 2452
catTas1r3        TGCAGATGGGCACCATCCTCCTCTGTGCCCTGGGTATCCTAGCCACCTTCCACCTGCCCA 2446
humanTAS1R3      TGCAGATGGGCGCCCTCCTGCTCTGTGTCCTGGGCATCCTGGCTGCCTTCCACCTGCCCA 2437
                       **    *     *   *              * *       *        ** *

mouseTas1r2      AGTGTTACATGATCCTTTTCTACCCGGAGCGCAACACTTCAGCTTATTTCAATAGCATGA 2503
ratTas1r2        AGTGTTACATGATCCTTTTCTACCCGGAGCGCAACACCTCAGCCTATTTCAATAGCATGA 2503
humanTAS1R2      AGTGCTACATGATCCTCTTCTACCCGGAGCGCAACACGCCCGCCTACTTCAACAGCATGA 2491
mouseTas1r1      AATGCTACGTGATTCTCTGCCGTCCAGAACTCAACAACACAGAACACTTTCAGGCCTCCA 2494
ratTas1r1        AGTGCTATGTGATTCTCTGCCGTCCAGAACTCAACAATACAGAACACTTTCAGGCCTCCA 2488
humanTAS1R1      AGTGCTACGTGATCCTCTGCCGCCCAGACCTCAACAGCACAGAGCACTTCCAGGCCTCCA 2491
mouseTas1r3      AGTGCTATGTGCTTCTTTGGCTGCCAAAGCTCAACACCCAGGAGTTCTTCCTGGGAAGGA 2512
ratTas1r3        AATGCTATGTACTTCTGTGGCTGCCAGAGCTCAACACCCAGGAGTTCTTCCTGGGAAGGA 2512
catTas1r3        AGTGCTACCTGCTGCTGCAGCGGCCGGAGCTCAACACCCCTGAGTTCTTCCTGGAAGACA 2506
humanTAS1R3      GGTGTTACCTGCTCATGCGGCAGCCAGGGCTCAACACCCCCGAGTTCTTCCTGGGAGGGG 2497
                   ** **  *  *  *       **    * *****     *     **

mouseTas1r2      --TTCAGGGCTACACGATGAGGAAGAGCTAG----------------------------- 2532
ratTas1r2        --TCCAGGGCTACACCATGAGGAAGAGC-------------------------------- 2529
humanTAS1R2      --TCCAGGGCTACACCATGAGGAGGGACTAG----------------------------- 2520
mouseTas1r1      --TCCAGGACTACACGAGGCGCTGCGGCACTACCTGA----------------------- 2529
ratTas1r1        --TCCAGGACTACACGAGGCGCTGCGGCACTACC-------------------------- 2520
humanTAS1R1      --TTCAGGACTACACGAGGCGCTGCGGCTCCACCTGA----------------------- 2526
mouseTas1r3      ATGCCAAGAAAGCAGCAGATGAGAAC-AGTGGCGGTGGTGAGGCAGCTCAGGGACACAAT 2571
ratTas1r3        GCCCCAAGGAAGCATCAGATGGGAAT-AGTGGTAGTAGTGAGGCAACTCGGGGACACAGT 2571
catTas1r3        ATGCCA---GAGCACAGGGCAGCAGTTGGGGGCAGGGGAGGGGAGAATCGGGGCAAAAAC 2563
humanTAS1R3      GCCCTGGGGATGCCCAAGGCCAGAAT----GACGGGAACACAGGAAATCAGGGGAAACAT 2553
                             *

mouseTas1r2      ------
ratTas1r2        ------
humanTAS1R2      ------
mouseTas1r1      ------
ratTas1r1        ------
humanTAS1R1      ------
mouseTas1r3      GAATGA 2577
ratTas1r3        GAATGA 2577
catTas1r3        AAGTGA 2569
humanTAS1R3      GAGTGA 2559
```

Figure 2A
CLUSTAL W (1.82) multiple amin acid sequence alignment f T1Rs:

```
humanT1R3       MLGPAVLGLS----LWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSR-- 54
catT1R3         MPGLALLGLTALLGLTALLDHGEGATSCLSQQLRMQGDYVLGGLFPLGSAEGTGLGDG-- 58
mouseT1R3       MPALAIMGLS----LAAFLELGMGASLCLSQQFKAQGDYILGGLFPLGSTEEATLNQR-- 54
ratT1R3         MPGLAILGLS----LAAFLELGMGSSLCLSQQFKAQGDYILGGLFPLGTTEEATLNQR-- 54
mouseT1R1       MLFWAAHLLLSLQLAVAYCWAFSCQRTESSPGFSLPGDFLLAGLFSLHADCLQVRHR--P 58
ratT1R1         MLFWAAHLLLSLQL--VYCWAFSCQRTESSPGFSLPGDFLLAGLFSLHGDCLQVRHR--P 56
humanT1R1       MLLCTARLVG-LQLLISCCWAFACHSTESSPDFTLPGDYLLAGLFPLHSGCLQVRHR--P 57
mouseT1R2       --MGPQARTLHLLFLLLHALPKPVMLVGNSD-FHLAGDYLLGGLFTLHANVKSVSHLSYL 57
ratT1R2         --MGPQARTLCLLSLLLHVLPKPGKLVENSD-FHLAGDYLLGGLFTLHANVKSISHLSYL 57
humanT1R2       --MGPRAKTICSLFFLLWVLAEP---AENSD-FYLPGDYLLGGLFSLHANMKGIVHLNFL 54
                    .                        *  :   **::*.***.*

                ♥    ♥ (possible functional amino acid substitution)
humanT1R3       TRPSSPVCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSL 114
catT1R3         LQPNATVCTRFSSLGLLWALAVKMAVEEINNGSALLPGLHLGYDLFDTCSEPMVAMKPSL 118
mouseT1R3       TQPNSIPCNRFSPLGLFLAMAMKMAVEEINNGSALLPGLRLGYDLFDTCSEPVVTMKSSL 114
ratT1R3         TQPNGILCTRFSPLGLFLAMAMKMAVEEINNGSALLPGLRLGYDLFDTCSEPVVTMKPSL 114
mouseT1R1       LVTSCDRSDSFNGHGYHLFQAMRFTVEEINNSTALLPNITLGYELYDVCSE-SSNVYATL 117
ratT1R1         LVTSCDRPDSFNGHGYHLFQAMRFTVEEINNSSALLPNITLGYELYDVCSE-SANVYATL 115
humanT1R1       EVTLCDRSCSFNEHGYHLFQAMRLGVEEINNSTALLPNITLGYQLYDVCSD-SANVYATL 116
mouseT1R2       QVPKCN-EYNMKVLGYNLMQAMRFAVEEINNCSSLLPGVLLGYEMVDVCYL-SNNIQPGL 115
ratT1R2         QVPKCN-EFTMKVLGYNLMQAMRFAVEEINNCSSLLPGVLLGYEMVDVCYL-SNNIHPGL 115
humanT1R2       QVPMCK-EYEVKVIGYNLMQAMRFAVEEINNDSSLLPGVLLGYEIVDVCYI-SNNVQPVL 112
                  .       ..  *     *::: ****** : ***.: ***:: *.*      : . *

humanT1R3       MFLAKAGSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELL 174
catT1R3         VFMAKAGSCSIAAYCNYTQYQPRVLAVIGPHSSELALVTGKFFSFFLVPQVSYGASTDRL 178
mouseT1R3       MFLAKVGSQSIAAYCNYTQYQPRVLAVIGPHSSELALITGKFFSFFLMPQVSYSASMDRL 174
ratT1R3         MFMAKVGSQSIAAYCNYTQYQPRVLAVIGPHSSELALITGKFFSFFLMPQVSYSASMDRL 174
mouseT1R1       RVLAQQGTGHLEMQRDLRNHSSKVVALIGPDNTDHAVTTAALLSPFLMPLVSYEASSVIL 177
ratT1R1         RVLALQGPRHIEIQKDLRNHSSKVVAFIGPDNTDHAVTTAALLGPFLMPLVSYEASSVVL 175
humanT1R1       RVLSLPGQHHIELQGDLLHYSPTVLAVIGPDSTNRAATTAALLSPFLVPMISYAASSETL 176
mouseT1R2       YFLSQI-DDFLPILKDYSQYRPQVVAVIGPDNSESAITVSNILSYFLVPQVTYSAITDKL 174
ratT1R2         YFLAQD-DDLLPILKDYSQYMPHVVAVIGPDNSESAITVSNILSHFLIPQITYSAISDKL 174
humanT1R2       YFLAHE-DNLLPIQEDYSNYISRVVAVIGPDNSESVMTVANFLSLFLLPQITYSAISDEL 171
                 .::      :    :  :: . *:*.***..:: .  .. ::. **:* ::* *    *

humanT1R3       SARETFPSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAAR- 233
catT1R3         SNREIFPSFFRTVPSDQVQVAAMVELLEELGWNWVAAVGSDDEYGRQGLSLFSGLASAR- 237
mouseT1R3       SDRETFPSFFRTVPSDRVQLQAVVTLLQNFSWNWVAALGSDDDYGREGLSIFSSLANAR- 233
ratT1R3         SDRETFPSFFRTVPSDRVQLQAVVTLLQNFSWNWVAALGSDDDYGREGLSIFSGLANSR- 233
mouseT1R1       SGKRKFPSFLRTIPSDKYQVEVIVRLLQSFGWVWISLVGSYGDYGQLGVQALEELATPR- 236
ratT1R1         SAKRKFPSFLRTVPSDRHQVEVMVQLLQSFGWVWISLIGSYGDYGQLGVQALEELAVPR- 234
humanT1R1       SVKRQYPSFLRTIPNDKYQVETMVLLLQKFGWTWISLVGSSDDYGQLGVQALENQATGQ- 235
mouseT1R2       RDKRRFPAMLRTVPSATHHIEAMVQLMVHFQWNWIVVLVSDDDYGRENSHLLSQRLTNTG 234
ratT1R2         RDKRHFPSMLRTVPSATHHIEAMVQLMVHFQWNWIVVLVSDDDYGRENSHLLSQRLTKTS 234
humanT1R2       RDKVRFPALLRTTPSADHHVEAMVQLMLHFRWNWIIVLVSSDTYGRDNGQLLGERVARR- 230
                  :  :*:::** *.   :: . . *:  : * *:  : * . **: .   :

humanT1R3       GICIAHEGLVPLPRADDSR----LGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSIS 289
catT1R3         GICIAHEGLVPLP-PGSLR----LGALQGLLRQVNQSSVQVVVLFSSAHAARTLFSYSIR 292
mouseT1R3       GICIAHEGLVPQHDTSGQQ----LGKVLDVLRQVNQSKVQVVVLFASARAVYSLFSYSIH 289
ratT1R3         GICIAHEGLVPQHDTSGQQ----LGKVVDVLRQVNQSKVQVVVLFASARAVYSLFSYSIL 289
mouseT1R1       GICVAFKDVVPLSAQAG------DPRMQRMMLRLARARTTVVVVFSNRHLAGVFFRSVVL 290
ratT1R1         GICVAFKDIVPFSARVG------DPRMQSMMQHLAQARTTVVVVFSNRHLARVFFRSVVL 288
humanT1R1       GICIAFKDIMPFSAQVG------DERMQCLMRHLAQAGATVVVVFSSRQLARVFFESVVL 289
mouseT1R2       DICIAFQEVLPVPEPNQAVRPEEQDQLDNILDKLRRTSARVVVIFSPELSLHNFFREVLR 294
ratT1R2         DICIAFQEVLPIPESSQVMRSEEQRQLDNILDKLRRTSARVVVVFSPELSLYSFFHEVLR 294
humanT1R2       DICIAFQETLPTLQPNQNMTSEERQRLVTIVDKLQQSTARVVVVFSPDLTLYHFFNEVLR 290
                .**:*.:  :*               :  :: :: :: . **::*:      :*    :
```

Figure 2B

```
humanT1R3       SRLSPKVWVASEAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPA 349
catT1R3         CKLSPKVWVASEAWLTSDLVMTLPGMPGVGTVLGFLQQGAPMPEFPSYVRTRLALAADPA 352
mouseT1R3       HGLSPKVWVASESWLTSDLVMTLPNIARVGTVLGFLQRGALLPEFSHYVETHLALAADPA 349
ratT1R3         HDLSPKVWVASESWLTSDLVMTLPNIARVGTVLGFLQRGALLPEFSHYVETRLALAADPT 349
mouseT1R1       ANLTGKVWIASEDWAISTYITNVPGIQGIGTVLGVAIQQRQVPGLKEFEESYVQAVMGAP 350
ratT1R1         ANLTGKVWVASEDWAISTYITSVTGIQGIGTVLGVAVQQRQVPGLKEFEESYVRAVTAAP 348
humanT1R1       TNLTGKVWVASEAWALSRHITGVPGIQRIGMVLGVAIQKRAVPGLKAFEEAYARADKKAP 349
mouseT1R2       WNFTGFVWIASESWAIDPVLHNLTELRHTGTFLGVTIQRVSIPGFSQFRVRHDKPEYPMP 354
ratT1R2         WNFTGFVWIASESWAIDPVLHNLTELRHTGTFLGVTIQRVSIPGFSQFRVRRDKPGYPVP 354
humanT1R2       QNFTGAVWIASESWAIDPVLHNLTELGHLGTFLGITIQSVPIPGFSEFREWGPQAGPPPL 350
                  ::  **:*** *  .  :  :. :   * .**.  :   :  :  :

humanT1R3       FCSALGEREQGLEEDVVGQRCPQCDCITLQNVS--------AGLNHHQTFSVYAAVYSVA 401
catT1R3         FCASLDAEQPGLEEHVVGPRCPQCDHVTLENLS--------AGLLHHQTFAAYAAVYGVA 404
mouseT1R3       FCASLN-AELDLEEHVMGQRCPRCDDIMLQNLSSGLLQNLSAGQLHHQIFATYAAVYSVA 408
ratT1R3         FCASLK-AELDLEERVMGPRCSQCDYIMLQNLSSGLMQNLSAGQLHHQIFATYAAVYSVA 408
mouseT1R1       RTCPEG------SWCGTNQLCRECHAFTTWNMP------ELGAFSMSAAYNVYEAVYAVA 398
ratT1R1         SACPEG------SWCSTNQLCRECHTFTTRNMP------TLGAFSMSAAYRVYEAVYAVA 396
humanT1R1       RPCHKG------SWCSSNQLCRECQAFMAHTMP------KLKAFSMSSAYNAYRAVYAVA 397
mouseT1R2       NETSLR------TTC--NQDCDACMNITESFNN------VLMLSGERVVYSVYSAVYAVA 400
ratT1R2         NTTNLR------TTC--NQDCDACLNTTKSFNN------ILILSGERVVYSVYSAVYAVA 400
humanT1R2       SRTSQS------YTC--NQECDNCLNATLSFNT------ILRLSGERVVYSVYSAVYAVA 396
                                 .  *  *                         : .* ***.**

humanT1R3       QALHNTLQCNASGCPAQDPVKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVW 461
catT1R3         QALHNTLRCNASGCPRREPVRPWQLLENMYNVSFRARGLALQFDASGNVNVDYDLKLWVW 464
mouseT1R3       QALHNTLQCNVSHCHVSEHVLPWQLLENMYNMSFHARDLTLQFDAEGNVDMEYDLKMWVW 468
ratT1R3         QALHNTLQCNVSHCHTSEPVQPWQLLENMYNMSFRARDLTLQFDAKGSVDMEYDLKMWVW 468
mouseT1R1       HGLHQLLGCTSGTCA-RGPVYPWQLLQQIYKVNFLLHKKTVAFDDKGDPLGYYDIIAWDW 457
ratT1R1         HGLHQLLGCTSEICS-RGPVYPWQLLQQIYKVNFLLHENTVAFDDNGDTLGYYDIIAWDW 455
humanT1R1       HGLHQLLGCASGACS-RGRVYPWQLLEQIHKVHFLLHKDTVAFNDNRDPLSSYNIIAWDW 456
mouseT1R2       HTLHRLLHCNQVRCT-KQIVYPWQLLREIWHVNFTLLGNQLFFDEQGDMPMLLDIIQWQW 459
ratT1R2         HALHRLLGCNRVRCT-KQKVYPWQLLREIWHVNFTLLGNRLFFDQQGDMPMLLDIIQWQW 459
humanT1R2       HALHSLLGCDKSTCT-KRVVYPWQLLEEIWKVNFTLLDHQIFFDPQGDVALHLEIVQWQW 455
                : **  * *    *     * *****.:: :: *      : *: . .     ::  * *

humanT1R3       QGSVPRLHDVGRFNG---SLRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCC 518
catT1R3         QDPTPELRTVGTFKG---RLELWRSQMCWHTPGKQQPVSQCSRQCKEGQVRRVKGFHSCC 521
mouseT1R3       QSPTPVLHTVGTFNG---TLQLQQSKMYWP--GNQVPVSQCSRQCKDGQVRRVKGFHSCC 523
ratT1R3         QSPTPVLHTVGTFNG---TLQLQHSKMYWP--GNQVPVSQCSRQCKDGQVRRVKGFHSCC 523
mouseT1R1       NGPEWTFEVIGSASLSPVHLDINKTKIQWHGKNNQVPVSVCTRDCLEGHHRLVMGSHHCC 517
ratT1R1         NGPEWTFEIIGSASLSPVHLDINKTKIQWHGKNNQVPVSVCTTDCLAGHHRVVVGSHHCC 515
humanT1R1       NGPKWTFTVLGSSTWSPVQLNINETKIQWHGKDNQVPKSVCSSDCLEGHQRVVTGFHHCC 516
mouseT1R2       GLSQNPFQSIASYSPTETRLTY-ISNVSWYTPNNTVPISMCSKSCQPGQMKKPIGLHPCC 518
ratT1R2         DLSQNPFQSIASYSPTSKRLTY-INNVSWYTPNNTVPVSMCSKSCQPGQMKKSVGLHPCC 518
humanT1R2       DRSQNPFQSVASYYPLQRQLKN-IQDISWHTVNNTIPMSMCSKRCQSGQKKKPVGIHVCC 514
                  .   :  :.        *     .: *   .:  * * *:  *  *: :   * * **

humanT1R3       YDCVDCEAGSYRQ-NPDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLS 577
catT1R3         YNCVDCKAGSYQR-NPDDLLCTQCDQDQWSPDRSTRCFARKPMFLAWGEPAVLLLLALLA 580
mouseT1R3       YDCVDCKAGSYRK-HPDDFTCTPCNQDQWSPEKSTACLPRRPKFLAWGEPVVLSLLLLLC 582
ratT1R3         YDCVDCKAGSYRK-HPDDFTCTPCGKDQWSPEKSTTCLPRRPKFLAWGEPAVLSLLLLLC 582
mouseT1R1       FECMPCEAGTFLN-TSELHTCQPCGTEEWAPEGSSACFSRTVEFLGWHEPISLVLLAANT 576
ratT1R1         FECVPCEAGTFLN-MSELHICQPCGTEEWAPKESTTCFPRTVEFLAWHEPISLVLIAANT 574
humanT1R1       FECVPCGAGTFLN-KSDLYRCQPCGKEEWAPEGSQTCFPRTVVFLALREHTSWVLLAANT 575
mouseT1R2       FECVDCPPGTYLNRSVDEFNCLSCPGSMWSYKNNIACFKRRLAFLEWHEVPTIVVTILAA 578
ratT1R2         FECLDCMPGTYLNRSADEFNCLSCPGSMWSYKNDITCFQRRPTFLEWHEVPTIVVAILAA 578
humanT1R2       FECIDCLPGTFLNHTEDEYECQACPNNEWSYQSETSCFKRQLVFLEWHEAPTIAVALLAA 574
                ::*: * .*:: .   :   *  *  . *: . .  *: *   **   *     :

humanT1R3       LALGLVLAALGLFVHHRDSPLVQASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQ 637
catT1R3         LALGLALAALGLFLWHSDSPLVQASGGPRACFGLACLGLVCLSVLLFPGQPGPASCLAQQ 640
mouseT1R3       LVLGLALAALGLSVHHWDSPLVQASGGSQFCFGLICLGLFCLSVLLFPGRPSSASCLAQQ 642
ratT1R3         LVLGLTLAALGLFVHYWDSPLVQASGGSLFCFGLICLGLFCLSVLLFPGRPRSASCLAQQ 642
mouseT1R1       LLLLLLIGTAGLFAWRLHTPVVRSAGGRLCFLMLGSLVAGSCSLYSFFGKPTVPACLLRQ 636
ratT1R1         LLLLLLVGTAGLFAWHFHTPVVRSAGGRLCFLMLGSLVAGSCSFYSFFGEPTVPACLLRQ 634
humanT1R1       LLLLLLLGTAGLFAWHLDTPVVRSAGGRLCFLMLGSLAAGSGSLYGFFGEPTRPACLLRQ 635
mouseT1R2       LGFISTLAILLIFWRHFQTPMVRSAGGPMCFLMLVPLLLAFGMVPVYVGPPTVFSCFCRQ 638
ratT1R2         LGFFSTLAILFIFWRHFQTPMVRSAGGPMCFLMLVPLLLAFGMVPVYVGPPTVFSCFCRQ 638
humanT1R2       LGFLSTLAILVIFWRHFQTPIVRSAGGPMCFLMLTLLLVAYMVVPVYVGPPKVSTCLCRQ 634
                * :   :.   :     .:*:*:::**    : *  *      .  : * *    *: :*
```

Figure 2C

```
humanT1R3        PLSHLPLTGCLSTLFLQAAEIFVESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCT 697
catT1R3          PLFHLPLTGCLSTFFLQAAEIFVGSELPPSWAEKMRGRLRGPWAWLVVLLAMLAEAALCA 700
mouseT1R3        PMAHLPLTGCLSTLFLQAAETFVESELPLSWANWLCSYLRGLWAWLVVLLATFVEAALCA 702
ratT1R3          PMAHLPLTGCLSTLFLQAAEIFVESELPLSWANWLCSYLRGPWAWLVVLLATLVEAALCA 702
mouseT1R1        PLFSLGFAIFLSCLTIRSFQLVIIFKFSTKVPTFYHTWAQNHGAGIFVIVSSTVHLFLCL 696
ratT1R1          PLFSLGFAIFLSCLTIRSFQLVIIFKFSTKVPTFYRTWAQNHGAGLFVIVSSTVHLLICL 694
humanT1R1        ALFALGFTIFLSCLTVRSFQLIIIFKFSTKVPTFYHAWVQNHGAGLFVMISSAAQLLICL 695
mouseT1R2        AFFTVCFSVCLSCITVRSFQIVCVFKMARRLPSAYGFWMRYHGPYVFVAFITAVKVALVA 698
ratT1R2          AFFTVCFSICLSCITVRSFQIVCVFKMARRLPSAYSFWMRYHGPYVFVAFITAIKVALVV 698
humanT1R2        ALFPLCFTICISCIAVRSFQIVCAFKMASRFPRAYSYWVRYQGPYVSMAFITVLKMVIVV 694
                 .:  : ::  :* : ::: : .   ::.   .       :   . : : .    .  :

humanT1R3        WYLVAFPPEVV-TDWHMLPTEALVHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPG 756
catT1R3          WYLVAFPPEVV-TDWRVLPTEALVHCHVHSWISFGLVHATNAMLAFLCFLGTFLVQSRPG 759
mouseT1R3        WYLIAFPPEVV-TDWSVLPTEVLEHCHVRSWVSLGLVHITNAMLAFLCFLGTFLVQSQPG 761
ratT1R3          WYLMAFPPEVV-TDWQVLPTEVLEHCRMRSWVSLGLVHITNAVLAFLCFLGTFLVQSQPG 761
mouseT1R1        TWLAMWTPRPT-REYQRFPHLVILECTEVNSVGFLVAFAHNILLSISTFVCSYLGKELPE 755
ratT1R1          TWLVMWTPRPT-REYQRFPHLVILECTEVNSVGFLLAFTHNILLSISTFVCSYLGKELPE 753
humanT1R1        TWLVVWTPLPA-REYQRFPHLVMLECTETNSLGFILAFLYNGLLSISAFACSYLGKDLPE 754
mouseT1R2        GNMLATTINPIGRTDPDDPNIIILSCHPNYRNGLLFNTSMDLLLSVLGFSFAYVGKELPT 758
ratT1R2          GNMLATTINPIGRTDPDDPNIMILSCHPNYRNGLLFNTSMDLLLSVLGFSFAYMGKELPT 758
humanT1R2        IGMLATGLSPTTRTDPDDPKITIVSCNPNYRNSLLFNTSLDLLLSVVGFSFAYMGKELPT 754
                   :               *   :  *      .: .    :  *:.  *  ::: :. *

humanT1R3        RYNRARGLTFAMLAYFITWVSFVPLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYL 816
catT1R3          RYNGARGLTFAMLAYFITWISFVPLFANVHVAYQPAVQMGTILLCALGILATFHLPKCYL 819
mouseT1R3        RYNRARGLTFAMLAYFITWVSFVPLLANVQVAYQPAVQMGAILVCALGILVTFHLPKCYV 821
ratT1R3          RYNRARGLTFAMLAYFIIWVSFVPLLANVQVAYQPAVQMGAILFCALGILATFHLPKCYV 821
mouseT1R1        NYNEAKCVTFSLLLHFVSWIAFFTMSSIYQGSYLPAVNVLAGLATLSGGFSGYFLPKCYV 815
ratT1R1          NYNEAKCVTFSLLLNFVSWIAFFTMASIYQGSYLPAVNVLAGLTTLSGGFSGYFLPKCYV 813
humanT1R1        NYNEAKCVTFSLLFNFVSWIAFFTTASVYDGKYLPAANMMAGLSSLSSGFGGYFLPKCYV 814
mouseT1R2        NYNEAKFITLSMTFSFTSSISLCTFMSVHDGVLVTIMDLLVTVLNFLAIGLGYFGPKCYM 818
ratT1R2          NYNEAKFITLSMTFSFTSSISLCTFMSVHDGVLVTIMDLLVTVLNFLAIGLGYFGPKCYM 818
humanT1R2        NYNEAKFITLSMTFYFTSSVSLCTFMSAYSGVLVTIVDLLVTVLNLLAISLGYFGPKCYM 814
                 .** *: :*:::   *   ::: .  :       .  :: . :    .    :. *:**:

humanT1R3        LMRQPGLNTPEFFLGG---GPGDAQGQNDGNTGNQGKHE------- 852
catT1R3          LLQRPELNTPEFFLEDNARAQGSSWGQGRGESGQKQVTPDPVTSPQ 865
mouseT1R3        LLWLPKLNTQEFFLGRN--AKKAADENSGGGEAAQGHNE------- 858
ratT1R3          LLWLPELNTQEFFLGRS--PKEASDGNSGSSEATRGHSE------- 858
mouseT1R1        ILCRPELNNTEHFQASIQDYTRRCGTT------------------- 842
ratT1R1          ILCRPELNNTEHFQASIQDYTRRCGTT------------------- 840
humanT1R1        ILCRPDLNSTEHFQASIQDYTRRCGST------------------- 841
mouseT1R2        ILFYPERNTSAYFNSMIQGYTMRKS--------------------- 843
ratT1R2          ILFYPERNTSAYFNSMIQGYTMRKS--------------------- 843
humanT1R2        ILFYPERNTPAYFNSMIQGYTMRRD--------------------- 839
                 ::  *  *.  .*
```

Figure 3

Phylogenetic Tree of T1Rs:

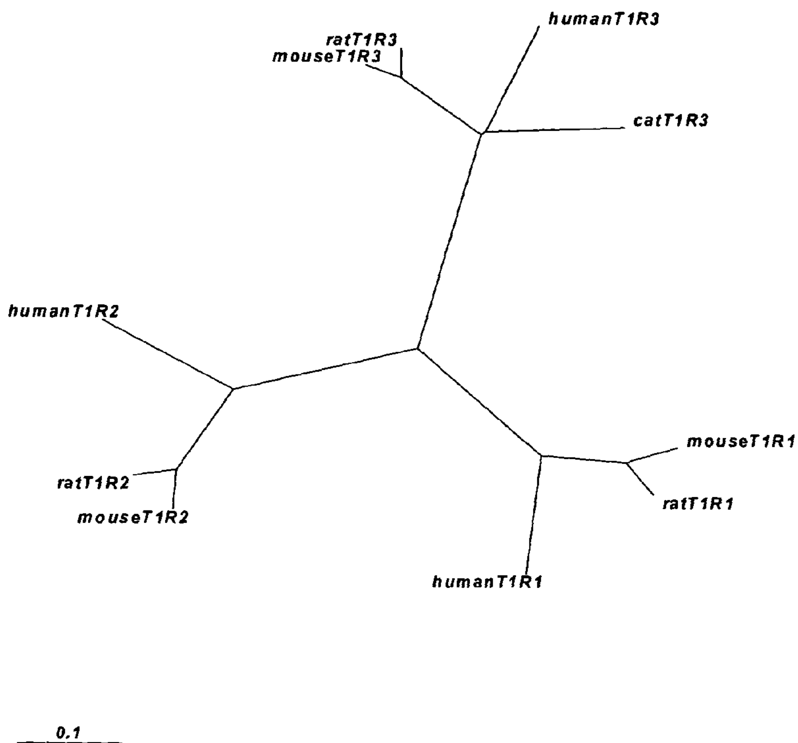
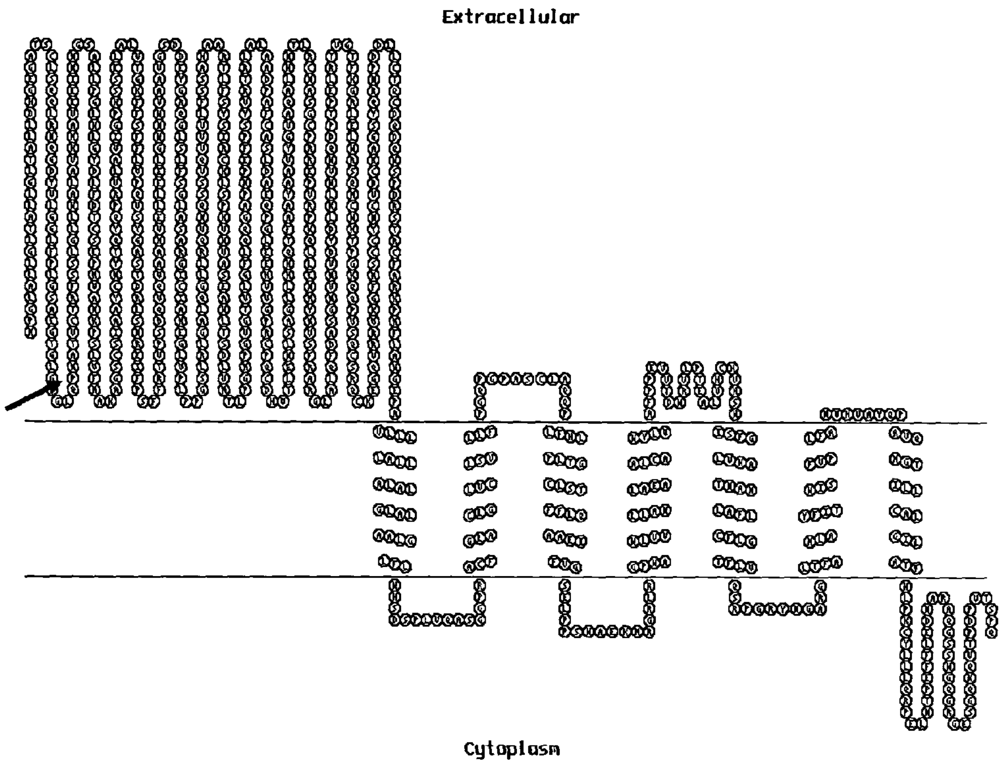


Figure 4. Predicted conformation of the 7TM T1R3 protein sequ nc from cat. Arrow points to region of possibl functional amino acid substitution.



SEQUENCE LISTING

```
<110>  Li, Xia
       Li, Weihua
       Reed, Danielle
       Bachmanov, Alexander
       Brand, Joseph

<120>  Taste Receptor Of The T1R Family From Domestic Cat

<130>  MON3042

<160>  58

<170>  PatentIn version 3.2

<210>  1
<211>  2569
<212>  DNA
<213>  Felis catus

<400>  1
atgcccggcc tcgctctcct gggcctcacg gctctcctgg gcctcacggc tctcttggac      60

cacggggagg gcgcaacgtc ctgcttgtca cagcagctca ggatgcaggg ggactatgtg     120

ctgggtgggc tcttccctct gggctctgcc gagggtacag gtcttggcga cgggctgcag     180

cccaatgcca ccgtgtgcac caggttctcg tctctgggcc tgctctgggc gctggccgtg     240

aagatggcgg tggaggagat caacaacggg tcggccctgc tgcccgggct gcacctgggc     300

tatgacctct ttgacacgtg ttcagagccc atggtggcca tgaagcccag cctcgtgttc     360

atggccaaag caggcagctg cagcattgcc gcctactgca attacacaca gtaccagccc     420

cgcgtgctgg ccgtcatcgg gccccactcg tctgagctcg ccctcgtcac cggcaagttc     480

ttcagcttct tccttgtgcc tcaggtcagc tacggcgcca gcaccgaccg gctgagcaac     540

cgggagatct tcccgtcctt cttccgcacg gtgcccagcg accaggtgca ggtggcggcc     600

atggtggagc tgctggagga gctcggctgg aactgggtgg cggcggtggg tagtgacgac     660

gagtatggcc ggcagggcct gagcctcttc tccggcctgg ccagcgccag ggcatctgc      720

atcgcgcatg agggcctggt gccactgccg ccaggcagcc tgcggctggg cgccctacag     780

ggcctgctgc gccaggtgaa ccagagcagc gtgcaggtgg tggtgctgtt ctcctccgcc     840

cacgcggccc gcaccctctt cagctacagc atccgctgca agctctcacc caaggtgtgg     900

gtggccagcg aggcctggct gacctcagac ctggtcatga cgctgcccgg catgcctggg     960

gtgggcaccg tgctgggctt cctgcagcag ggcgccccga tgccggagtt cccatcctac    1020

gtgcggaccc gcctggccct ggccgctgac cctgccttct gcgcctcgct ggacgctgaa    1080

cagccaggcc tggaggagca cgtggtgggg ccacgctgcc cccaatgtga ccacgtcacg    1140

ctagagaacc tatctgcggg gctgctgcac caccagacct tcgctgccta cgcggctgtg    1200

tatggcgtgg cccaagccct tcacaacaca ctgcgctgca atgcctcggg ctgccccagg    1260

cgggagcctg tgcggccctg gcagctccta gagaacatgt acaacgtgag cttccgtgct    1320

cgcggcctgg cactgcagtt cgacgccagc gggaacgtga acgtggatta cgacctgaaa    1380

ctgtgggtgt ggcaggaccc gacgcccgag ctgcgcaccg taggcacctt caagggccgc    1440

ctggagctct ggcgctctca gatgtgctgg cacacgccgg ggaagcagca gcccgtgtcc    1500
```

cagtgctccc ggcagtgcaa ggaaggccag gtgcgccgcg tgaagggctt ccactcttgc 1560

tgttacaact gcgtggactg caaggcgggc agttatcagc gcaacccaga tgacctcctc 1620

tgcacccagt gtgaccagga ccagtggtcc ccagaccgga gcacacgctg cttcgcccgc 1680

aagcccatgt tcctggcatg gggggagcca gctgtgctgc tactgctcgc gctgctggct 1740

ctggcgctgg gcctggcgct ggcagccctg gggctcttcc tctggcactc ggacagcccg 1800

ctggttcagg cctcaggtgg gccacgggcc tgctttggcc tggcttgcct gggcctggtc 1860

tgcctcagtg tcctcctgtt ccctggccag ccaggccctg ccagctgcct ggcccagcag 1920

ccactgttcc acctcccact cactggctgc ctgagcacgt ttttcctgca agcggccgag 1980

atatttgtgg ggtcggagct gccaccaagc tgggctgaga agatgcgtgg ccgcctgcgg 2040

gggccctggg cctggctggt ggtgctgctt gctatgctgg cagaagccgc attgtgtgcc 2100

tggtacctgg tagccttccc gccagaggtg gtgacggact ggcgggtact gcccacagag 2160

gcgctggtgc actgccacgt gcactcctgg atcagcttcg gcctggtgca tgccactaac 2220

gccatgctgg ccttcctctg cttcctgggc actttcctgg tgcagagccg gccaggccgc 2280

tacaatggtg cccgcggcct cacctttgcc atgctggcct acttcatcac ctggatctcc 2340

tttgtgcccc tctttgccaa tgtgcacgtg gcctaccagc ctgccgtgca gatgggcacc 2400

atcctcctct gtgccctggg tatcctagcc accttccacc tgcccaagtg ctacctgctg 2460

ctgcagcggc cggagctcaa cacccctgag ttcttcctgg aagacaatgc cagagcacag 2520

ggcagcagtt gggggcaggg gaggggagaa tcggggcaaa aacaagtga 2569

<210> 2
<211> 865
<212> PRT
<213> Felis catus

<400> 2

Met Pro Gly Leu Ala Leu Leu Gly Leu Thr Ala Leu Leu Gly Leu Thr
1 5 10 15

Ala Leu Leu Asp His Gly Glu Gly Ala Thr Ser Cys Leu Ser Gln Gln
20 25 30

Leu Arg Met Gln Gly Asp Tyr Val Leu Gly Gly Leu Phe Pro Leu Gly
35 40 45

Ser Ala Glu Gly Thr Gly Leu Gly Asp Gly Leu Gln Pro Asn Ala Thr
50 55 60

Val Cys Thr Arg Phe Ser Ser Leu Gly Leu Leu Trp Ala Leu Ala Val
65 70 75 80

Lys Met Ala Val Glu Glu Ile Asn Asn Gly Ser Ala Leu Leu Pro Gly
85 90 95

Leu His Leu Gly Tyr Asp Leu Phe Asp Thr Cys Ser Glu Pro Met Val
100 105 110

Ala Met Lys Pro Ser Leu Val Phe Met Ala Lys Ala Gly Ser Cys Ser
115 120 125

Ile Ala Ala Tyr Cys Asn Tyr Thr Gln Tyr Gln Pro Arg Val Leu Ala
130 135 140

Val Ile Gly Pro His Ser Ser Glu Leu Ala Leu Val Thr Gly Lys Phe
145 150 155 160

Phe Ser Phe Phe Leu Val Pro Gln Val Ser Tyr Gly Ala Ser Thr Asp
165 170 175

Arg Leu Ser Asn Arg Glu Ile Phe Pro Ser Phe Phe Arg Thr Val Pro
180 185 190

Ser Asp Gln Val Gln Val Ala Ala Met Val Glu Leu Leu Glu Glu Leu
195 200 205

Gly Trp Asn Trp Val Ala Ala Val Gly Ser Asp Asp Glu Tyr Gly Arg
210 215 220

Gln Gly Leu Ser Leu Phe Ser Gly Leu Ala Ser Ala Arg Gly Ile Cys
225 230 235 240

Ile Ala His Glu Gly Leu Val Pro Leu Pro Pro Gly Ser Leu Arg Leu
245 250 255

Gly Ala Leu Gln Gly Leu Leu Arg Gln Val Asn Gln Ser Ser Val Gln
260 265 270

Val Val Val Leu Phe Ser Ser Ala His Ala Ala Arg Thr Leu Phe Ser
275 280 285

Tyr Ser Ile Arg Cys Lys Leu Ser Pro Lys Val Trp Val Ala Ser Glu
290 295 300

Ala Trp Leu Thr Ser Asp Leu Val Met Thr Leu Pro Gly Met Pro Gly
305 310 315 320

Val Gly Thr Val Leu Gly Phe Leu Gln Gln Gly Ala Pro Met Pro Glu
325 330 335

Phe Pro Ser Tyr Val Arg Thr Arg Leu Ala Leu Ala Ala Asp Pro Ala
340 345 350

Phe Cys Ala Ser Leu Asp Ala Glu Gln Pro Gly Leu Glu Glu His Val
355 360 365

Val Gly Pro Arg Cys Pro Gln Cys Asp His Val Thr Leu Glu Asn Leu
370 375 380

Ser Ala Gly Leu Leu His His Gln Thr Phe Ala Ala Tyr Ala Ala Val

385 390 395 400

Tyr Gly Val Ala Gln Ala Leu His Asn Thr Leu Arg Cys Asn Ala Ser
405 410 415

Gly Cys Pro Arg Arg Glu Pro Val Arg Pro Trp Gln Leu Leu Glu Asn
420 425 430

Met Tyr Asn Val Ser Phe Arg Ala Arg Gly Leu Ala Leu Gln Phe Asp
435 440 445

Ala Ser Gly Asn Val Asn Val Asp Tyr Asp Leu Lys Leu Trp Val Trp
450 455 460

Gln Asp Pro Thr Pro Glu Leu Arg Thr Val Gly Thr Phe Lys Gly Arg
465 470 475 480

Leu Glu Leu Trp Arg Ser Gln Met Cys Trp His Thr Pro Gly Lys Gln
485 490 495

Gln Pro Val Ser Gln Cys Ser Arg Gln Cys Lys Glu Gly Gln Val Arg
500 505 510

Arg Val Lys Gly Phe His Ser Cys Cys Tyr Asn Cys Val Asp Cys Lys
515 520 525

Ala Gly Ser Tyr Gln Arg Asn Pro Asp Asp Leu Leu Cys Thr Gln Cys
530 535 540

Asp Gln Asp Gln Trp Ser Pro Asp Arg Ser Thr Arg Cys Phe Ala Arg
545 550 555 560

Lys Pro Met Phe Leu Ala Trp Gly Glu Pro Ala Val Leu Leu Leu Leu
565 570 575

Ala Leu Leu Ala Leu Ala Leu Gly Leu Ala Leu Ala Ala Leu Gly Leu
580 585 590

Phe Leu Trp His Ser Asp Ser Pro Leu Val Gln Ala Ser Gly Gly Pro
595 600 605

Arg Ala Cys Phe Gly Leu Ala Cys Leu Gly Leu Val Cys Leu Ser Val
610 615 620

Leu Leu Phe Pro Gly Gln Pro Gly Pro Ala Ser Cys Leu Ala Gln Gln
625 630 635 640

Pro Leu Phe His Leu Pro Leu Thr Gly Cys Leu Ser Thr Phe Phe Leu
645 650 655

Gln Ala Ala Glu Ile Phe Val Gly Ser Glu Leu Pro Pro Ser Trp Ala
660 665 670

Glu Lys Met Arg Gly Arg Leu Arg Gly Pro Trp Ala Trp Leu Val Val
675 680 685

Leu Leu Ala Met Leu Ala Glu Ala Ala Leu Cys Ala Trp Tyr Leu Val
690 695 700

Ala Phe Pro Pro Glu Val Val Thr Asp Trp Arg Val Leu Pro Thr Glu
705 710 715 720

Ala Leu Val His Cys His Val His Ser Trp Ile Ser Phe Gly Leu Val
725 730 735

His Ala Thr Asn Ala Met Leu Ala Phe Leu Cys Phe Leu Gly Thr Phe
740 745 750

Leu Val Gln Ser Arg Pro Gly Arg Tyr Asn Gly Ala Arg Gly Leu Thr
755 760 765

Phe Ala Met Leu Ala Tyr Phe Ile Thr Trp Ile Ser Phe Val Pro Leu
770 775 780

Phe Ala Asn Val His Val Ala Tyr Gln Pro Ala Val Gln Met Gly Thr
785 790 795 800

Ile Leu Leu Cys Ala Leu Gly Ile Leu Ala Thr Phe His Leu Pro Lys
805 810 815

Cys Tyr Leu Leu Leu Gln Arg Pro Glu Leu Asn Thr Pro Glu Phe Phe
820 825 830

Leu Glu Asp Asn Ala Arg Ala Gln Gly Ser Ser Trp Gly Gln Gly Arg
835 840 845

Gly Glu Ser Gly Gln Lys Gln Val Thr Pro Asp Pro Val Thr Ser Pro
850 855 860

Gln
865

<210> 3
<211> 2532
<212> DNA
<213> Mus musculus

<400> 3
atgggacccc aggcgaggac actccatttg ctgtttctcc tgctgcatgc tctgcctaag 60

ccagtcatgc tggtagggaa ctccgacttt cacctggctg gggactacct cctgggtggc 120

ctctttaccc tccatgccaa cgtgaagagc gtctctcacc tcagctacct gcaggtgccc 180

aagtgcaatg agtacaacat gaaggtcttg ggctacaacc tcatgcaggc catgcgattc 240

gccgtggagg aaatcaacaa ctgtagctct ctgctgcccg gcgtgctgct cggctacgag 300

atggtggatg tctgctacct ctccaacaat atccagcctg ggctctactt cctgtcacag 360

atagatgact tcctgcccat cctcaaagac tacagccagt acaggcccca agtggtggcc 420

gtcattggcc cagacaactc tgagtccgcc atcaccgtgt ccaacattct ctcctacttc 480

ctcgtgccac aggtcacata tagcgccatc accgacaagc tgcgagacaa gcggcgcttc 540

cctgccatgc tgcgcactgt gcccagcgcc acccaccaca tcgaggccat ggtgcaactg 600

atggttcact tccagtggaa ctggatcgtg gtgctggtga gcgatgacga ttatggccga 660

gagaacagcc acctgctgag ccagcgtctg accaacactg gcgatatctg cattgccttc 720

caggaggttc tgcctgtacc agaacccaac caggccgtga ggcctgagga gcaggaccaa 780

ctggacaaca tcctggacaa gctgcggcgg acctcggcgc gtgtggtggt gatattctcg 840

ccagagctga gcctgcacaa cttcttccgc gaggtgctgc gctggaactt cacaggcttt 900

gtgtggattg cctctgagtc ctgggccatc gaccctgttc tacacaacct cacagagctg 960

cgccacacgg gcactttcct gggcgtcacc atccagaggg tgtccatccc tggcttcagc 1020

cagttccgag tgcgccacga caagccagag tatcccatgc ctaacgagac cagcctgagg 1080

actacctgta accaggactg tgacgcctgc atgaacatca ccgagtcctt taacaacgtt 1140

ctcatgcttt cgggggagcg tgtggtctac agtgtgtact cggccgtcta cgcggtagcc 1200

cacaccctcc acagactcct ccactgcaac caggtccgct gcaccaagca aatcgtctat 1260

ccatggcagc tactcaggga gatctggcat gtcaacttca cgctcctggg caaccagctc 1320

ttcttcgacg aacaagggga catgccgatg ctcctggaca tcatccagtg gcaatggggc 1380

ctgagccaga accccttcca aagcatcgcc tcctactccc ccaccgagac gaggctgacc 1440

tacattagca atgtgtcctg gtacaccccc aacaacacgg tccccatatc catgtgttct 1500

aagagttgcc agcctgggca aatgaaaaaa cccataggcc tccacccgtg ctgcttcgag 1560

tgtgtggact gtccgccggg cacctacctc aaccgatcag tagatgagtt taactgtctg 1620

tcctgcccgg gttccatgtg gtcttacaag aacaacatcg cttgcttcaa gcggcggctg 1680

gccttcctgg agtggcacga agtgcccact atcgtggtga ccatcctggc cgccctgggc 1740

ttcatcagta cgctggccat tctgctcatc ttctggagac atttccagac gcccatggtg 1800

cgctcggcgg gcggccccat gtgcttcctg atgctggtgc ccctgctgct ggcgttcggg 1860

atggtccccg tgtatgtggg cccccccacg gtcttctcct gtttctgccg ccaggctttc 1920

ttcaccgttt gcttctccgt ctgcctctcc tgcatcacgg tgcgctcctt ccagattgtg 1980

tgcgtcttca agatggccag acgcctgcca agcgcctacg gtttctggat gcgttaccac 2040

gggccctacg tctttgtggc cttcatcacg gccgtcaagg tggccctggt ggcaggcaac 2100

atgctggcca ccaccatcaa ccccattggc cggaccgacc ccgatgaccc caatatcata 2160

atcctctcct gccaccctaa ctaccgcaac gggctactct tcaacaccag catggacttg 2220

ctgctgtccg tgctgggttt cagcttcgcg tacgtgggca aggaactgcc caccaactac 2280

aacgaagcca agttcatcac cctcagcatg accttctcct tcacctcctc catctccctc 2340

tgcacgttca tgtctgtcca cgatggcgtg ctggtcacca tcatggatct cctggtcact 2400

gtgctcaact ttctggccat cggcttgggg tactttggcc ccaagtgtta catgatcctt 2460

ttctacccgg agcgcaacac ttcagcttat ttcaatagca tgattcaggg ctacacgatg 2520

aggaagagct ag 2532

<210> 4
<211> 2529
<212> DNA
<213> Rattus rattus

<400> 4
atgggtcccc aggcaaggac actctgcttg ctgtctctcc tgctgcatgt tctgcctaag 60

ccaggcaagc tggtagagaa ctctgacttc cacctggccg gggactacct cctgggtggc 120

ctctttaccc tccatgccaa cgtgaagagc atctcccacc tcagctacct gcaggtgccc 180

aagtgcaatg agttcaccat gaaggtgttg ggctacaacc tcatgcaggc catgcgtttc 240

gctgtggagg agatcaacaa ctgtagctcc ctgctacccg gcgtgctgct cggctacgag 300

atggtggatg tctgttacct ctccaacaat atccaccctg ggctctactt cctggcacag 360

gacgacgacc tcctgcccat cctcaaagac tacagccagt acatgcccca cgtggtggct 420

gtcattggcc ccgacaactc tgagtccgcc attaccgtgt ccaacattct ctctcatttc 480

ctcatcccac agatcacata cagcgccatc tccgacaagc tgcgggacaa gcggcacttc 540

cctagcatgc tacgcacagt gcccagcgcc acccaccaca tcgaggccat ggtgcagctg 600

atggttcact tccaatggaa ctggattgtg gtgctggtga gcgacgacga ttacggccgc 660

gagaacagcc acctgttgag ccagcgtctg accaaaacga gcgacatctg cattgccttc 720

caggaggttc tgcccatacc tgagtccagc caggtcatga ggtccgagga gcagagacaa 780

ctggacaaca tcctggacaa gctgcggcgg acctcggcgc gcgtcgtggt ggtgttctcg 840

cccgagctga gcctgtatag cttctttcac gaggtgctcc gctggaactt cacgggtttt 900

gtgtggatcg cctctgagtc ctgggctatc gacccagttc tgcataacct cacggagctg 960

cgccacacgg gtacttttct gggcgtcacc atccagaggg tgtccatccc tggcttcagt 1020

cagttccgag tgcgccgtga caagccaggg tatcccgtgc ctaacacgac caacctgcgg 1080

acgacctgca accaggactg tgacgcctgc ttgaacacca ccaagtcctt caacaacatc 1140

cttatacttt cgggggagcg cgtggtctac agcgtgtact cggcagttta cgcggtggcc 1200

catgccctcc acagactcct cggctgtaac cgggtccgct gcaccaagca aaaggtctac 1260

ccgtggcagc tactcaggga gatctggcac gtcaacttca cgctcctggg taaccggctc 1320

ttctttgacc aacaagggga catgccgatg ctcttggaca tcatccagtg gcagtgggac 1380

ctgagccaga atcccttcca aagcatcgcc tcctattctc ccaccagcaa gaggctaacc 1440

tacattaaca atgtgtcctg gtacaccccc aacaacacgg tccctgtctc catgtgttcc 1500

aagagctgcc agccagggca aatgaaaaag tctgtgggcc tccacccttg ttgcttcgag 1560

tgcttggatt gtatgccagg cacctacctc aaccgctcag cagatgagtt taactgtctg 1620

tcctgcccgg gttccatgtg gtcctacaag aacgacatca cttgcttcca gcggcggcct 1680

accttcctgg agtggcacga agtgcccacc atcgtggtgg ccatactggc tgccctgggc 1740

ttcttcagta cactggccat tcttttcatc ttctggagac atttccagac acccatggtg 1800

cgctcggccg gtggccccat gtgcttcctg atgctcgtgc ccctgctgct ggcgtttggg 1860

atggtgcccg tgtatgtggg gcccccacg gtcttctcat gcttctgccg acaggctttc 1920

ttcaccgtct gcttctccat ctgcctatcc tgcatcaccg tgcgctcctt ccagatcgtg 1980

tgtgtcttca agatggccag acgcctgcca agtgcctaca gtttttggat gcgttaccac 2040

gggccctatg tcttcgtggc cttcatcacg gccatcaagg tggccctggt ggtgggcaac 2100

atgctggcca ccaccatcaa ccccattggc cggaccgacc cggatgaccc caacatcatg 2160

atcctctcgt gccaccctaa ctaccgcaac gggctactgt tcaacaccag catggacttg 2220

ctgctgtctg tgctgggttt cagcttcgct tacatgggca aggagctgcc caccaactac 2280

aacgaagcca agttcatcac tctcagcatg accttctcct tcacctcctc catctccctc 2340

tgcaccttca tgtctgtgca cgacggcgtg ctggtcacca tcatggacct cctggtcact 2400

gtgctcaact tcctggccat cggcttggga tactttggcc ccaagtgtta catgatcctt 2460

ttctacccgg agcgcaacac ctcagcctat ttcaatagca tgatccaggg ctacaccatg 2520

aggaagagc 2529

<210> 5
<211> 2520
<212> DNA
<213> Homo sapiens

<400> 5
atggggccca gggcaaagac catctgctcc ctgttcttcc tcctatgggt cctggctgag 60

ccggctgaga actcggactt ctacctgcct ggggattacc tcctgggtgg cctcttctcc 120

ctccatgcca acatgaaggg cattgttcac cttaacttcc tgcaggtgcc catgtgcaag 180

gagtatgaag tgaaggtgat aggctacaac ctcatgcagg ccatgcgctt cgcggtggag 240

gagatcaaca atgacagcag cctgctgcct ggtgtgctgc tgggctatga gatcgtggat 300

gtgtgctaca tctccaacaa tgtccagccg gtgctctact tcctggcaca cgaggacaac 360

ctccttccca tccaagagga ctacagtaac tacatttccc gtgtggtggc tgtcattggc 420

cctgacaact ccgagtctgt catgactgtg gccaacttcc tctccctatt tctccttcca 480

cagatcacct acagcgccat cagcgatgag ctgcgagaca aggtgcgctt cccggctttg 540

ctgcgtacca cacccagcgc cgaccaccac gtcgaggcca tggtgcagct gatgctgcac 600

ttccgctgga actggatcat tgtgctggtg agcagcgaca cctatggccg cgacaatggc 660

cagctgcttg gcgagcgcgt ggcccggcgc gacatctgca tcgccttcca ggagacgctg 720

cccacactgc agcccaacca gaacatgacg tcagaggagc gccagcgcct ggtgaccatt 780

gtggacaagc tgcagcagag cacagcgcgc gtcgtggtcg tgttctcgcc cgacctgacc 840

ctgtaccact tcttcaatga ggtgctgcgc cagaacttca cgggcgccgt gtggatcgcc 900

tccgagtcct gggccatcga cccggtcctg cacaacctca cggagctggg ccacttgggc 960

accttcctgg gcatcaccat ccagagcgtg cccatcccgg gcttcagtga gttccgcgag 1020

tggggcccac aggctgggcc gccaccctc agcaggacca gccagagcta tacctgcaac 1080

caggagtgcg acaactgcct gaacgccacc ttgtccttca acaccattct caggctctct 1140

ggggagcgtg tcgtctacag cgtgtactct gcggtctatg ctgtggccca tgccctgcac 1200

agcctcctcg gctgtgacaa aagcacctgc accaagaggg tggtctaccc ctggcagctg 1260

cttgaggaga tctggaaggt caacttcact ctcctggacc accaaatctt cttcgacccg 1320

caaggggacg tggctctgca cttggagatt gtccagtggc aatgggaccg gagccagaat 1380

cccttccaga gcgtcgcctc ctactacccc ctgcagcgac agctgaagaa catccaagac 1440

atctcctggc acaccgtcaa caacacgatc cctatgtcca tgtgttccaa gaggtgccag 1500

tcagggcaaa agaagaagcc tgtgggcatc cacgtctgct gcttcgagtg catcgactgc 1560

cttcccggca ccttcctcaa ccacactgaa gatgaatatg aatgccaggc ctgcccgaat 1620

aacgagtggt cctaccagag tgagacctcc tgcttcaagc ggcagctggt cttcctggaa 1680

tggcatgagg cacccaccat cgctgtggcc ctgctggccg ccctgggctt cctcagcacc 1740

ctggccatcc tggtgatatt ctggaggcac ttccagacac ccatagttcg ctcggctggg 1800

ggccccatgt gcttcctgat gctgacactg ctgctggtgg catacatggt ggtcccggtg 1860

tacgtggggc cgcccaaggt ctccacctgc ctctgccgcc aggccctctt tcccctctgc 1920

ttcacaattt gcatctcctg tatcgccgtg cgttctttcc agatcgtctg cgccttcaag 1980

atggccagcc gcttcccacg cgcctacagc tactgggtcc gctaccaggg gccctacgtc 2040

tctatggcat ttatcacggt actcaaaatg gtcattgtgg taattggcat gctggccacg 2100

ggcctcagtc ccaccacccg tactgacccc gatgacccca agatcacaat tgtctcctgt 2160

aaccccaact accgcaacag cctgctgttc aacaccagcc tggacctgct gctctcagtg 2220

gtgggtttca gcttcgccta catgggcaaa gagctgccca ccaactacaa cgaggccaag 2280

ttcatcaccc tcagcatgac cttctatttc acctcatccg tctccctctg caccttcatg 2340

tctgcctaca gcggggtgct ggtcaccatc gtggacctct tggtcactgt gctcaacctc 2400

ctggccatca gcctgggcta cttcggcccc aagtgctaca tgatcctctt ctacccggag 2460

cgcaacacgc ccgcctactt caacagcatg atccagggct acaccatgag gagggactag 2520

<210> 6
<211> 2529
<212> DNA
<213> Mus musculus

<400> 6
atgcttttct gggcagctca cctgctgctc agcctgcagc tggccgttgc ttactgctgg 60

gctttcagct gccaaaggac agaatcctct ccaggtttca gcctccctgg ggacttcctc 120

ctggcaggcc tgttctccct ccatgctgac tgtctgcagg tgagacacag acctctggtg 180

acaagttgtg acaggtctga cagcttcaac ggccatggct atcacctctt ccaagccatg 240

cggttcaccg ttgaggagat aaacaactcc acagctctgc ttcccaacat caccctgggg 300

tatgaactgt atgacgtgtg ctcagagtct tccaatgtct atgccaccct gagggtgctc 360

gcccagcaag ggacaggcca cctagagatg cagagagatc ttcgcaacca ctcctccaag 420

gtggtggcac tcattgggcc tgataacact gaccacgctg tcaccactgc tgccctgctg 480

agcccttttc tgatgcccct ggtcagctat gaggcgagca gcgtgatcct cagtgggaag 540

cgcaagttcc cgtccttctt gcgcaccatc cccagcgata agtaccaggt ggaagtcata 600

gtgcggctgc tgcagagctt cggctgggtc tggatctcgc tcgttggcag ctatggtgac 660

tacgggcagc tgggcgtaca ggcgctggag gagctggcca ctccacgggg catctgcgtc 720

gccttcaagg acgtggtgcc tctctccgcc caggcgggtg acccaaggat gcagcgcatg 780

atgctgcgtc tggctcgagc caggaccacc gtggtcgtgg tcttctctaa ccggcacctg 840

gctggagtgt tcttcaggtc tgtggtgctg gccaacctga ctggcaaagt gtggatcgcc 900

tccgaagact gggccatctc cacgtacatc accaatgtgc ccgggatcca gggcattggg 960

acggtgctgg gggtggccat ccagcagaga caagtccctg gcctgaagga gtttgaagag 1020

tcctatgtcc aggcagtgat gggtgctccc agaacttgcc cagaggggtc ctggtgcggc 1080

actaaccagc tgtgcaggga gtgtcacgct ttcacgacat ggaacatgcc cgagcttgga 1140

gccttctcca tgagcgctgc ctacaatgtg tatgaggctg tgtatgctgt ggcccacggc 1200

ctccaccagc tcctgggatg tacctctggg acctgtgcca gaggcccagt ctacccctgg 1260

cagcttcttc agcagatcta caaggtgaat ttccttctac ataagaagac tgtagcattc 1320

gatgacaagg gggaccctct aggttattat gacatcatcg cctgggactg gaatggacct 1380

gaatggacct ttgaggtcat tggttctgcc tcactgtctc cagttcatct agacataaat 1440

aagacaaaaa tccagtggca cgggaagaac aatcaggtgc ctgtgtcagt gtgtaccagg 1500

gactgtctcg aagggcacca caggttggtc atgggttccc accactgctg cttcgagtgc 1560

atgccctgtg aagctgggac atttctcaac acgagtgagc ttcacacctg ccagccttgt 1620

ggaacagaag aatgggcccc tgaggggagc tcagcctgct tctcacgcac cgtggagttc 1680

ttggggtggc atgaacccat ctctttggtg ctattagcag ctaacacgct attgctgctg 1740

ctgctgattg ggactgctgg cctgtttgcc tggcgtcttc acacgcctgt tgtgaggtca 1800

gctgggggta ggctgtgctt cctcatgctg ggttccttgg tagctgggag ttgcagcctc 1860

tacagcttct tcgggaagcc cacggtgccc gcgtgcttgc tgcgtcagcc cctcttttct 1920

ctcgggtttg ccattttcct ctcctgtctg acaatccgct ccttccaact ggtcatcatc 1980

ttcaagtttt ctaccaaggt acccacattc taccacactt gggcccaaaa ccatggtgcc 2040

ggaatattcg tcattgtcag ctccacggtc catttgttcc tctgtctcac gtggcttgca 2100

atgtggaccc cacggcccac cagggagtac cagcgcttcc cccatctggt gattcttgag 2160

tgcacagagg tcaactctgt gggcttcctg gtggctttcg cacacaacat cctcctctcc 2220

atcagcacct ttgtctgcag ctacctgggt aaggaactgc cggagaacta taacgaagcc 2280

aaatgtgtca ccttcagcct gctcctccac ttcgtatcct ggatcgcttt cttcaccatg 2340

tccagcattt accagggcag ctacctaccc gcggtcaatg tgctggcagg gctggccact 2400

ctgagtggcg gcttcagcgg ctatttcctc cctaaatgct acgtgattct ctgccgtcca 2460

gaactcaaca acacagaaca ctttcaggcc tccatccagg actacacgag gcgctgcggc 2520

actacctga 2529

<210> 7
<211> 2520
<212> DNA
<213> Rattus rattus

<400> 7
atgctcttct gggctgctca cctgctgctc agcctgcagt tggtctactg ctgggctttc 60

agctgccaaa ggacagagtc ctctccaggc ttcagccttc ctggggactt cctccttgca 120

ggtctgttct ccctccatgg tgactgtctg caggtgagac acagacctct ggtgacaagt 180

tgtgacaggc ccgacagctt caacggccat ggctaccacc tcttccaagc catgcggttc 240

actgttgagg agataaacaa ctcctcggcc ctgcttccca acatcaccct ggggtatgag 300

ctgtacgacg tgtgctcaga atctgccaat gtgtatgcca ccctgagggt gcttgccctg 360

caagggcccc gccacataga gatacagaaa gaccttcgca accactcctc caaggtggtg 420

gccttcatcg ggcctgacaa cactgaccac gctgtcacta ccgctgcctt gctgggtcct 480

ttcctgatgc ccctggtcag ctatgaggca agcagcgtgg tactcagtgc caagcgcaag 540

ttcccgtctt tccttcgtac cgtccccagt gaccggcacc aggtggaggt catggtgcag 600

ctgctgcaga gttttgggtg ggtgtggatc tcgctcattg gcagctacgg tgattacggg 660

cagctgggtg tgcaggcgct ggaggagctg gccgtgcccc ggggcatctg cgtcgccttc 720

aaggacatcg tgcctttctc tgcccgggtg ggtgacccga ggatgcagag catgatgcag 780

catctggctc aggccaggac caccgtggtt gtggtcttct ctaaccggca cctggctaga 840

gtgttcttca ggtccgtggt gctggccaac ctgactggca aagtgtgggt cgcctcagaa 900

gactgggcca tctccacgta catcaccagc gtgactggga tccaaggcat tgggacggtg 960

ctcggtgtgg ccgtccagca gagacaagtc cctgggctga aggagtttga ggagtcttat 1020

gtcagggctg taacagctgc tcccagcgct tgcccggagg ggtcctggtg cagcactaac 1080

cagctgtgcc gggagtgcca cacgttcacg actcgtaaca tgcccacgct tggagccttc 1140

tccatgagtg ccgcctacag agtgtatgag gctgtgtacg ctgtggccca cggcctccac 1200

cagctcctgg gatgtacttc tgagatctgt tccagaggcc cagtctaccc ctggcagctt 1260

cttcagcaga tctacaaggt gaattttctt ctacatgaga atactgtggc atttgatgac 1320

aacgggggaca ctctaggtta ctacgacatc atcgcctggg actggaatgg acctgaatgg 1380

acctttgaga tcattggctc tgcctcactg tctccagttc atctggacat aaataagaca 1440

aaaatccagt ggcacgggaa gaacaatcag gtgcctgtgt cagtgtgtac cacggactgt 1500

ctggcagggc accacagggt ggttgtgggt tcccaccact gctgctttga gtgtgtgccc 1560

tgcgaagctg ggacctttct caacatgagt gagcttcaca tctgccagcc ttgtggaaca 1620

gaagaatggg cacccaagga gagcactact tgcttcccac gcacggtgga gttcttggct 1680

tggcatgaac ccatctcttt ggtgctaata gcagctaaca cgctattgct gctgctgctg 1740

gttgggactg ctggcctgtt tgcctggcat tttcacacac ctgtagtgag gtcagctggg 1800

ggtaggctgt gcttcctcat gctgggttcc ctggtggccg gaagttgcag cttctatagc 1860

ttcttcgggg agcccacggt gcccgcgtgc ttgctgcgtc agcccctctt ttctctcggg 1920

tttgccatct tcctctcctg cctgacaatc cgctccttcc aactggtcat catcttcaag 1980

ttttctacca aggtgcccac attctaccgt acctgggccc aaaaccatgg tgcaggtcta 2040

ttcgtcattg tcagctccac ggtccatttg ctcatctgtc tcacatggct tgtaatgtgg 2100

accccacgac ccaccaggga ataccagcgc ttccccccatc tggtgattct cgagtgcaca 2160

gaggtcaact ctgtaggctt cctgttggct ttcacccaca acattctcct ctccatcagt 2220

accttcgtct gcagctacct gggtaaggaa ctgccagaga actataatga agccaaatgt 2280

gtcaccttca gcctgctcct caacttcgta tcctggatcg ccttcttcac catggccagc 2340

atttaccagg gcagctacct gcctgcggtc aatgtgctgg cagggctgac cacactgagc 2400

ggcggcttca gcggttactt cctcccccaag tgctatgtga ttctctgccg tccagaactc 2460

aacaatacag aacactttca ggcctccatc caggactaca cgaggcgctg cggcactacc 2520

<210> 8
<211> 2526
<212> DNA
<213> Homo sapiens

<400> 8
atgctgctct gcacggctcg cctggtcggc ctgcagcttc tcatttcctg ctgctgggcc 60

tttgcctgcc atagcacgga gtcttctcct gacttcaccc tccccggaga ttacctcctg 120

gcaggcctgt tccctctcca ttctggctgt ctgcaggtga ggcacagacc cgaggtgacc 180

ctgtgtgaca ggtcttgtag cttcaatgag catggctacc acctcttcca ggctatgcgg 240

cttggggttg aggagataaa caactccacg gccctgctgc ccaacatcac cctggggtac 300

cagctgtatg atgtgtgttc tgactctgcc aatgtgtatg ccacgctgag agtgctctcc 360

ctgccagggc aacaccacat agagctccaa ggagaccttc tccactattc ccctacggtg 420

ctggcagtga ttgggcctga cagcaccaac cgtgctgcca ccacagccgc cctgctgagc 480

cctttcctgg tgcccatgat tagctatgcg gccagcagcg agacgctcag cgtgaagcgg 540

cagtatccct ctttcctgcg caccatcccc aatgacaagt accaggtgga gaccatggtg 600

ctgctgctgc agaagttcgg gtggacctgg atctctctgg ttggcagcag tgacgactat 660

gggcagctag ggtgcaggc actggagaac caggccactg gtcaggggat ctgcattgct 720

ttcaaggaca tcatgccctt ctctgcccag gtgggcgatg agaggatgca gtgcctcatg 780

cgccacctgg cccaggccgg ggccaccgtc gtggttgttt tttccagccg gcagttggcc 840

agggtgtttt tcgagtccgt ggtgctgacc aacctgactg gcaaggtgtg ggtcgcctca 900

gaagcctggg ccctctccag gcacatcact ggggtgcccg ggatccagcg cattgggatg 960

gtgctgggcg tggccatcca gaagagggct gtccctggcc tgaaggcgtt tgaagaagcc 1020

tatgcccggg cagacaagaa ggcccctagg ccttgccaca agggctcctg gtgcagcagc 1080

aatcagctct gcagagaatg ccaagctttc atggcacaca cgatgcccaa gctcaaagcc 1140

ttctccatga gttctgccta caacgcatac cgggctgtgt atgcggtggc ccatggcctc 1200

caccagctcc tgggctgtgc ctctggagct tgttccaggg gccgagtcta cccctggcag 1260

cttttggagc agatccacaa ggtgcatttc cttctacaca aggacactgt ggcgtttaat 1320

gacaacagag atcccctcag tagctataac ataattgcct gggactggaa tggacccaag 1380

tggaccttca cggtcctcgg ttcctccaca tggtctccag ttcagctaaa cataaatgag 1440

accaaaatcc agtggcacgg aaaggacaac caggtgccta agtctgtgtg ttccagcgac 1500

tgtcttgaag ggcaccagcg agtggttacg ggtttccatc actgctgctt tgagtgtgtg 1560

ccctgtgggg ctgggacctt cctcaacaag agtgacctct acagatgcca gccttgtggg 1620

aaagaagagt gggcacctga gggaagccag acctgcttcc cgcgcactgt ggtgtttttg 1680

gctttgcgtg agcacacctc ttgggtgctg ctggcagcta acacgctgct gctgctgctg 1740

ctgcttggga ctgctggcct gtttgcctgg cacctagaca cccctgtggt gaggtcagca 1800

gggggccgcc tgtgctttct tatgctgggc tccctggcag caggtagtgg cagcctctat 1860

ggcttctttg gggaacccac aaggcctgcg tgcttgctac gccaggccct ctttgccctt 1920

ggtttcacca tcttcctgtc ctgcctgaca gttcgctcat tccaactaat catcatcttc 1980

aagttttcca ccaaggtacc tacattctac cacgcctggg tccaaaacca cggtgctggc 2040

ctgtttgtga tgatcagctc agcggcccag ctgcttatct gtctaacttg gctggtggtg 2100

tggaccccac tgcctgctag ggaataccag cgcttccccc atctggtgat gcttgagtgc 2160

acagagacca actccctggg cttcatactg gccttcctct acaatggcct cctctccatc 2220

agtgcctttg cctgcagcta cctgggtaag gacttgccag agaactacaa cgaggccaaa 2280

tgtgtcacct tcagcctgct cttcaacttc gtgtcctgga tcgccttctt caccacggcc 2340

agcgtctacg acggcaagta cctgcctgcg gccaacatga tggctgggct gagcagcctg 2400

agcagcggct tcggtgggta ttttctgcct aagtgctacg tgatcctctg ccgcccagac 2460

ctcaacagca cagagcactt ccaggcctcc attcaggact acacgaggcg ctgcggctcc 2520

acctga 2526

<210> 9
<211> 2577
<212> DNA
<213> Mus musculus

<400> 9
atgccagctt tggctatcat gggtctcagc ctggctgctt tcctggagct tgggatgggg 60

gcctctttgt gtctgtcaca gcaattcaag gcacaagggg actacatact gggcgggcta 120

tttcccctgg gctcaaccga ggaggccact ctcaaccaga gaacacaacc caacagcatc 180

ccgtgcaaca ggttctcacc ccttggtttg ttcctggcca tggctatgaa gatggctgtg 240

gaggagatca acaatggatc tgccttgctc cctgggctgc ggctgggcta tgacctattt 300

gacacatgct ccgagccagt ggtcaccatg aaatccagtc tcatgttcct ggccaaggtg 360

ggcagtcaaa gcattgctgc ctactgcaac tacacacagt accaaccccg tgtgctggct 420

gtcatcggcc cccactcatc agagcttgcc ctcattacag gcaagttctt cagcttcttc 480

ctcatgccac aggtcagcta tagtgccagc atggatcggc taagtgaccg ggaaacgttt 540

ccatccttct tccgcacagt gcccagtgac cgggtgcagc tgcaggcagt tgtgactctg 600

ttgcagaact tcagctggaa ctgggtggcc gccttaggga gtgatgatga ctatggccgg 660

gaaggtctga gcatcttttc tagtctggcc aatgcacgag gtatctgcat cgcacatgag 720

ggcctggtgc cacaacatga cactagtggc caacagttgg gcaaggtgct ggatgtacta 780

cgccaagtga accaaagtaa agtacaagtg gtggtgctgt ttgcctctgc ccgtgctgtc 840

tactcccttt ttagttacag catccatcat ggcctctcac ccaaggtatg ggtggccagt 900

gagtcttggc tgacatctga cctggtcatg acacttccca atattgcccg tgtgggcact 960

gtgcttgggt ttttgcagcg gggtgcccta ctgcctgaat tttcccatta tgtggagact 1020

caccttgccc tggccgctga cccagcattc tgtgcctcac tgaatgcgga gttggatctg 1080

gaggaacatg tgatggggca acgctgtcca cggtgtgacg acatcatgct gcagaaccta 1140

tcatctgggc tgttgcagaa cctatcagct gggcaattgc accaccaaat atttgcaacc 1200

tatgcagctg tgtacagtgt ggctcaagcc cttcacaaca ccctacagtg caatgtctca 1260

cattgccacg tatcagaaca tgttctaccc tggcagctcc tggagaacat gtacaatatg 1320

agtttccatg ctcgagactt gacactacag tttgatgctg aagggaatgt agacatggaa 1380

tatgacctga agatgtgggt gtggcagagc cctacacctg tattacatac tgtgggcacc 1440

ttcaacggca cccttcagct gcagcagtct aaaatgtact ggccaggcaa ccaggtgcca 1500

gtctcccagt gttcccgcca gtgcaaagat ggccaggttc gccgagtaaa gggctttcat 1560

tcctgctgct atgactgcgt ggactgcaag gcgggcagct accggaagca tccagatgac 1620

ttcacctgta ctccatgtaa ccaggaccag tggtccccag agaaaagcac agcctgctta 1680

cctcgcaggc ccaagtttct ggcttggggg gagccagttg tgctgtcact cctcctgctg 1740

ctttgcctgg tgctgggtct agcactggct gctctggggc tctctgtcca ccactgggac 1800

agccctcttg tccaggcctc aggtggctca cagttctgct ttggcctgat ctgcctaggc 1860

ctcttctgcc tcagtgtcct tctgttccca gggcggccaa gctctgccag ctgccttgca 1920

caacaaccaa tggctcacct ccctctcaca ggctgcctga gcacactctt cctgcaagca 1980

gctgagacct ttgtggagtc tgagctgcca ctgagctggg caaactggct atgcagctac 2040

cttcggggac tctgggcctg gctagtggta ctgttggcca cttttgtgga ggcagcacta 2100

tgtgcctggt atttgatcgc tttcccacca gaggtggtga cagactggtc agtgctgccc 2160

acagaggtac tggagcactg ccacgtgcgt tcctgggtca gcctgggctt ggtgcacatc 2220

accaatgcaa tgttagcttt cctctgcttt ctgggcactt tcctggtaca gagccagcct 2280

ggccgctaca accgtgcccg tggtctcacc ttcgccatgc tagcttattt catcacctgg 2340

gtctcttttg tgcccctcct ggccaatgtg caggtggcct accagccagc tgtgcagatg 2400

ggtgctatcc tagtctgtgc cctgggcatc ctggtcacct tccacctgcc caagtgctat 2460

gtgcttcttt ggctgccaaa gctcaacacc caggagttct tcctgggaag gaatgccaag 2520

aaagcagcag atgagaacag tggcggtggt gaggcagctc agggacacaa tgaatga 2577

<210> 10
<211> 2577
<212> DNA
<213> Rattus rattus

<400> 10
atgccgggtt tggctatctt gggcctcagt ctggctgctt tcctggagct tgggatgggg 60

tcctctttgt gtctgtcaca gcaattcaag gcacaagggg actatatatt gggtggacta 120

tttcccctgg gcacaactga ggaggccact ctcaaccaga gaacacagcc caacggcatc 180

ctatgtacca ggttctcgcc ccttggtttg ttcctggcca tggctatgaa gatggctgta 240
gaggagatca acaatggatc tgccttgctc cctgggctgc gactgggcta tgacctgttt 300
gacacatgct cagagccagt ggtcaccatg aagcccagcc tcatgttcat ggccaaggtg 360
ggaagtcaaa gcattgctgc ctactgcaac tacacacagt accaaccccg tgtgctggct 420
gtcattggtc cccactcatc agagcttgcc ctcattacag gcaagttctt cagcttcttc 480
ctcatgccac aggtcagcta tagtgccagc atggatcggc taagtgaccg ggaaacattt 540
ccatccttct tccgcacagt gcccagtgac cgggtgcagc tgcaggccgt tgtgacactg 600
ttgcagaatt tcagctggaa ctgggtggct gccttaggta gtgatgatga ctatggccgg 660
gaaggtctga gcatcttttc tggtctggcc aactcacgag gtatctgcat tgcacacgag 720
ggcctggtgc cacaacatga cactagtggc caacaattgg gcaaggtggt ggatgtgcta 780
cgccaagtga accaaagcaa agtacaggtg gtggtgctgt ttgcatctgc ccgtgctgtc 840
tactcccttt ttagctacag catccttcat gacctctcac ccaaggtatg ggtggccagt 900
gagtcctggc tgacctctga cctggtcatg acacttccca atattgcccg tgtgggcact 960
gttcttgggt ttctgcagcg cggtgcccta ctgcctgaat ttttccatta tgtggagact 1020
cgccttgccc tagctgctga cccaacattc tgtgcctccc tgaaagctga gttggatctg 1080
gaggagcgcg tgatggggcc acgctgttca caatgtgact acatcatgct acagaacctg 1140
tcatctgggc tgatgcagaa cctatcagct gggcagttgc accaccaaat atttgcaacc 1200
tatgcagctg tgtacagtgt ggctcaggcc cttcacaaca ccctgcagtg caatgtctca 1260
cattgccaca catcagagcc tgttcaaccc tggcagctcc tggagaacat gtacaatatg 1320
agtttccgtg ctcgagactt gacactgcag tttgatgcca aagggagtgt agacatggaa 1380
tatgacctga agatgtgggt gtggcagagc cctacacctg tactacatac tgtaggcacc 1440
ttcaacggca cccttcagct gcagcactcg aaaatgtatt ggccaggcaa ccaggtgcca 1500
gtctcccagt gctcccggca gtgcaaagat ggccaggtgc gcagagtaaa gggctttcat 1560
tcctgctgct atgactgtgt ggactgcaag gcagggagct accggaagca tccagatgac 1620
ttcacctgta ctccatgtgg caaggatcag tggtccccag aaaaaagcac aacctgctta 1680
cctcgcaggc ccaagtttct ggcttggggg gagccagctg tgctgtcact tctcctgctg 1740
ctttgcctgg tgctgggcct gacactggct gccctggggc tctttgtcca ctactgggac 1800
agccctcttg ttcaggcctc aggtgggtca ctgttctgct ttggcctgat ctgcctaggc 1860
ctcttctgcc tcagtgtcct tctgttccca ggacgaccac gctctgccag ctgccttgcc 1920
caacaaccaa tggctcacct ccctctcaca ggctgcctga gcacactctt cctgcaagca 1980
gccgagatct ttgtggagtc tgagctgcca ctgagttggg caaactggct ctgcagctac 2040
cttcggggcc cctgggcttg gctggtggta ctgctggcca ctcttgtgga ggctgcacta 2100
tgtgcctggt acttgatggc tttccctcca gaggtggtga cagattggca ggtgctgccc 2160
acggaggtac tggaacactg ccgcatgcgt tcctgggtca gcctgggctt ggtgcacatc 2220
accaatgcag tgttagcttt cctctgcttt ctgggcactt tcctggtaca gagccagcct 2280

ggtcgctata accgtgcccg tggcctcacc ttcgccatgc tagcttattt catcatctgg 2340

gtctcttttg tgcccctcct ggctaatgtg caggtggcct accagccagc tgtgcagatg 2400

ggtgctatct tattctgtgc cctgggcatc ctggccacct tccacctgcc caaatgctat 2460

gtacttctgt ggctgccaga gctcaacacc caggagttct tcctgggaag gagccccaag 2520

gaagcatcag atgggaatag tggtagtagt gaggcaactc ggggacacag tgaatga 2577

<210> 11
<211> 2559
<212> DNA
<213> Homo sapiens

<400> 11
atgctgggcc ctgctgtcct gggcctcagc ctctgggctc tcctgcaccc tgggacgggg 60

gccccattgt gcctgtcaca gcaacttagg atgaaggggg actacgtgct gggggggctg 120

ttcccctgg gcgaggccga ggaggctggc ctccgcagcc ggacacggcc cagcagccct 180

gtgtgcacca ggttctcctc aaacggcctg ctctgggcac tggccatgaa aatggccgtg 240

gaggagatca acaacaagtc ggatctgctg cccgggctgc gcctgggcta cgacctcttt 300

gatacgtgct cggagcctgt ggtggccatg aagcccagcc tcatgttcct ggccaaggca 360

ggcagccgcg acatcgccgc ctactgcaac tacacgcagt accagccccg tgtgctggct 420

gtcatcgggc cccactcgtc agagctcgcc atggtcaccg gcaagttctt cagcttcttc 480

ctcatgcccc aggtcagcta cggtgctagc atggagctgc tgagcgcccg ggagaccttc 540

ccctccttct tccgcaccgt gcccagcgac cgtgtgcagc tgacggccgc cgcggagctg 600

ctgcaggagt tcggctggaa ctgggtggcc gccctgggca gcgacgacga gtacggccgg 660

cagggcctga gcatcttctc ggccctggcc gcggcacgcg gcatctgcat cgcgcacgag 720

ggcctggtgc cgctgccccg tgccgatgac tcgcggctgg ggaaggtgca ggacgtcctg 780

caccaggtga accagagcag cgtgcaggtg gtgctgctgt tcgcctccgt gcacgccgcc 840

cacgccctct tcaactacag catcagcagc aggctctcgc ccaaggtgtg ggtggccagc 900

gaggcctggc tgacctctga cctggtcatg gggctgcccg gcatggccca gatgggcacg 960

gtgcttggct tcctccagag gggtgcccag ctgcacgagt tcccccagta cgtgaagacg 1020

cacctggccc tggccaccga cccggccttc tgctctgccc tgggcgagag ggagcagggt 1080

ctggaggagg acgtggtggg ccagcgctgc ccgcagtgtg actgcatcac gctgcagaac 1140

gtgagcgcag ggctaaatca ccaccagacg ttctctgtct acgcagctgt gtatagcgtg 1200

gcccaggccc tgcacaacac tcttcagtgc aacgcctcag gctgccccgc gcaggacccc 1260

gtgaagccct ggcagctcct ggagaacatg tacaacctga ccttccacgt gggcgggctg 1320

ccgctgcggt tcgacagcag cggaaacgtg gacatggagt acgacctgaa gctgtgggtg 1380

tggcagggct cagtgcccag gctccacgac gtgggcaggt tcaacggcag cctcaggaca 1440

gagcgcctga agatccgctg gcacacgtct gacaaccaga agcccgtgtc ccggtgctcg 1500

cggcagtgcc aggagggcca ggtgcgccgg gtcaaggggt tccactcctg ctgctacgac 1560

tgtgtggact gcgaggcggg cagctaccgg caaaacccag acgacatcgc ctgcaccttt 1620

tgtggccagg atgagtggtc cccggagcga agcacacgct gcttccgccg caggtctcgg 1680

ttcctggcat ggggcgagcc ggctgtgctg ctgctgctcc tgctgctgag cctggcgctg 1740

ggccttgtgc tggctgcttt ggggctgttc gttcaccatc gggacagccc actggttcag 1800

gcctcggggg ggcccctggc ctgctttggc ctggtgtgcc tgggcctggt ctgcctcagc 1860

gtcctcctgt tccctggcca gcccagccct gcccgatgcc tggcccagca gcccttgtcc 1920

cacctcccgc tcacgggctg cctgagcaca ctcttcctgc aggcggccga gatcttcgtg 1980

gagtcagaac tgcctctgag ctgggcagac cggctgagtg gctgcctgcg ggggccctgg 2040

gcctggctgg tggtgctgct ggccatgctg gtggaggtcg cactgtgcac ctggtacctg 2100

gtggccttcc cgccggaggt ggtgacggac tggcacatgc tgcccacgga ggcgctggtg 2160

cactgccgca cacgctcctg ggtcagcttc ggcctagcgc acgccaccaa tgccacgctg 2220

gcctttctct gcttcctggg cactttcctg gtgcggagcc agccgggccg ctacaaccgt 2280

gcccgtggcc tcacctttgc catgctggcc tacttcatca cctgggtctc ctttgtgccc 2340

ctcctggcca atgtgcaggt ggtcctcagg cccgccgtgc agatgggcgc cctcctgctc 2400

tgtgtcctgg gcatcctggc tgccttccac ctgcccaggt gttacctgct catgcggcag 2460

ccagggctca acacccccga gttcttcctg ggagggggcc ctggggatgc ccaaggccag 2520

aatgacggga acacaggaaa tcaggggaaa catgagtga 2559

<210> 12
<211> 852
<212> PRT
<213> Homo sapiens

<400> 12

Met Leu Gly Pro Ala Val Leu Gly Leu Ser Leu Trp Ala Leu Leu His
1 5 10 15

Pro Gly Thr Gly Ala Pro Leu Cys Leu Ser Gln Gln Leu Arg Met Lys
20 25 30

Gly Asp Tyr Val Leu Gly Gly Leu Phe Pro Leu Gly Glu Ala Glu Glu
35 40 45

Ala Gly Leu Arg Ser Arg Thr Arg Pro Ser Ser Pro Val Cys Thr Arg
50 55 60

Phe Ser Ser Asn Gly Leu Leu Trp Ala Leu Ala Met Lys Met Ala Val
65 70 75 80

Glu Glu Ile Asn Asn Lys Ser Asp Leu Leu Pro Gly Leu Arg Leu Gly
85 90 95

Tyr Asp Leu Phe Asp Thr Cys Ser Glu Pro Val Val Ala Met Lys Pro
100 105 110

Ser Leu Met Phe Leu Ala Lys Ala Gly Ser Arg Asp Ile Ala Ala Tyr
115 120 125

Cys Asn Tyr Thr Gln Tyr Gln Pro Arg Val Leu Ala Val Ile Gly Pro
    130   135   140

His Ser Ser Glu Leu Ala Met Val Thr Gly Lys Phe Phe Ser Phe Phe
145   150   155   160

Leu Met Pro Gln Val Ser Tyr Gly Ala Ser Met Glu Leu Leu Ser Ala
                165   170   175

Arg Glu Thr Phe Pro Ser Phe Phe Arg Thr Val Pro Ser Asp Arg Val
            180   185   190

Gln Leu Thr Ala Ala Ala Glu Leu Leu Gln Glu Phe Gly Trp Asn Trp
        195   200   205

Val Ala Ala Leu Gly Ser Asp Asp Glu Tyr Gly Arg Gln Gly Leu Ser
    210   215   220

Ile Phe Ser Ala Leu Ala Ala Ala Arg Gly Ile Cys Ile Ala His Glu
225   230   235   240

Gly Leu Val Pro Leu Pro Arg Ala Asp Asp Ser Arg Leu Gly Lys Val
                245   250   255

Gln Asp Val Leu His Gln Val Asn Gln Ser Ser Val Gln Val Val Leu
            260   265   270

Leu Phe Ala Ser Val His Ala Ala His Ala Leu Phe Asn Tyr Ser Ile
        275   280   285

Ser Ser Arg Leu Ser Pro Lys Val Trp Val Ala Ser Glu Ala Trp Leu
    290   295   300

Thr Ser Asp Leu Val Met Gly Leu Pro Gly Met Ala Gln Met Gly Thr
305   310   315   320

Val Leu Gly Phe Leu Gln Arg Gly Ala Gln Leu His Glu Phe Pro Gln
                325   330   335

Tyr Val Lys Thr His Leu Ala Leu Ala Thr Asp Pro Ala Phe Cys Ser
            340   345   350

Ala Leu Gly Glu Arg Glu Gln Gly Leu Glu Glu Asp Val Val Gly Gln
        355   360   365

Arg Cys Pro Gln Cys Asp Cys Ile Thr Leu Gln Asn Val Ser Ala Gly
    370   375   380

Leu Asn His His Gln Thr Phe Ser Val Tyr Ala Ala Val Tyr Ser Val
385   390   395   400

Ala Gln Ala Leu His Asn Thr Leu Gln Cys Asn Ala Ser Gly Cys Pro
                405   410   415

Ala Gln Asp Pro Val Lys Pro Trp Gln Leu Leu Glu Asn Met Tyr Asn
420 425 430

Leu Thr Phe His Val Gly Gly Leu Pro Leu Arg Phe Asp Ser Ser Gly
435 440 445

Asn Val Asp Met Glu Tyr Asp Leu Lys Leu Trp Val Trp Gln Gly Ser
450 455 460

Val Pro Arg Leu His Asp Val Gly Arg Phe Asn Gly Ser Leu Arg Thr
465 470 475 480

Glu Arg Leu Lys Ile Arg Trp His Thr Ser Asp Asn Gln Lys Pro Val
485 490 495

Ser Arg Cys Ser Arg Gln Cys Gln Glu Gly Gln Val Arg Arg Val Lys
500 505 510

Gly Phe His Ser Cys Cys Tyr Asp Cys Val Asp Cys Glu Ala Gly Ser
515 520 525

Tyr Arg Gln Asn Pro Asp Asp Ile Ala Cys Thr Phe Cys Gly Gln Asp
530 535 540

Glu Trp Ser Pro Glu Arg Ser Thr Arg Cys Phe Arg Arg Arg Ser Arg
545 550 555 560

Phe Leu Ala Trp Gly Glu Pro Ala Val Leu Leu Leu Leu Leu Leu Leu
565 570 575

Ser Leu Ala Leu Gly Leu Val Leu Ala Ala Leu Gly Leu Phe Val His
580 585 590

His Arg Asp Ser Pro Leu Val Gln Ala Ser Gly Gly Pro Leu Ala Cys
595 600 605

Phe Gly Leu Val Cys Leu Gly Leu Val Cys Leu Ser Val Leu Leu Phe
610 615 620

Pro Gly Gln Pro Ser Pro Ala Arg Cys Leu Ala Gln Gln Pro Leu Ser
625 630 635 640

His Leu Pro Leu Thr Gly Cys Leu Ser Thr Leu Phe Leu Gln Ala Ala
645 650 655

Glu Ile Phe Val Glu Ser Glu Leu Pro Leu Ser Trp Ala Asp Arg Leu
660 665 670

Ser Gly Cys Leu Arg Gly Pro Trp Ala Trp Leu Val Val Leu Leu Ala
675 680 685

Met Leu Val Glu Val Ala Leu Cys Thr Trp Tyr Leu Val Ala Phe Pro

690 695 700

Pro Glu Val Val Thr Asp Trp His Met Leu Pro Thr Glu Ala Leu Val
705 710 715 720

His Cys Arg Thr Arg Ser Trp Val Ser Phe Gly Leu Ala His Ala Thr
725 730 735

Asn Ala Thr Leu Ala Phe Leu Cys Phe Leu Gly Thr Phe Leu Val Arg
740 745 750

Ser Gln Pro Gly Arg Tyr Asn Arg Ala Arg Gly Leu Thr Phe Ala Met
755 760 765

Leu Ala Tyr Phe Ile Thr Trp Val Ser Phe Val Pro Leu Leu Ala Asn
770 775 780

Val Gln Val Val Leu Arg Pro Ala Val Gln Met Gly Ala Leu Leu Leu
785 790 795 800

Cys Val Leu Gly Ile Leu Ala Ala Phe His Leu Pro Arg Cys Tyr Leu
805 810 815

Leu Met Arg Gln Pro Gly Leu Asn Thr Pro Glu Phe Phe Leu Gly Gly
820 825 830

Gly Pro Gly Asp Ala Gln Gly Gln Asn Asp Gly Asn Thr Gly Asn Gln
835 840 845

Gly Lys His Glu
850

<210> 13
<211> 858
<212> PRT
<213> Mus musculus

<400> 13

Met Pro Ala Leu Ala Ile Met Gly Leu Ser Leu Ala Ala Phe Leu Glu
1 5 10 15

Leu Gly Met Gly Ala Ser Leu Cys Leu Ser Gln Gln Phe Lys Ala Gln
20 25 30

Gly Asp Tyr Ile Leu Gly Gly Leu Phe Pro Leu Gly Ser Thr Glu Glu
35 40 45

Ala Thr Leu Asn Gln Arg Thr Gln Pro Asn Ser Ile Pro Cys Asn Arg
50 55 60

Phe Ser Pro Leu Gly Leu Phe Leu Ala Met Ala Met Lys Met Ala Val
65 70 75 80

Glu Glu Ile Asn Asn Gly Ser Ala Leu Leu Pro Gly Leu Arg Leu Gly

85 90 95

Tyr Asp Leu Phe Asp Thr Cys Ser Glu Pro Val Val Thr Met Lys Ser
100 105 110

Ser Leu Met Phe Leu Ala Lys Val Gly Ser Gln Ser Ile Ala Ala Tyr
115 120 125

Cys Asn Tyr Thr Gln Tyr Gln Pro Arg Val Leu Ala Val Ile Gly Pro
130 135 140

His Ser Ser Glu Leu Ala Leu Ile Thr Gly Lys Phe Phe Ser Phe Phe
145 150 155 160

Leu Met Pro Gln Val Ser Tyr Ser Ala Ser Met Asp Arg Leu Ser Asp
165 170 175

Arg Glu Thr Phe Pro Ser Phe Phe Arg Thr Val Pro Ser Asp Arg Val
180 185 190

Gln Leu Gln Ala Val Val Thr Leu Leu Gln Asn Phe Ser Trp Asn Trp
195 200 205

Val Ala Ala Leu Gly Ser Asp Asp Asp Tyr Gly Arg Glu Gly Leu Ser
210 215 220

Ile Phe Ser Ser Leu Ala Asn Ala Arg Gly Ile Cys Ile Ala His Glu
225 230 235 240

Gly Leu Val Pro Gln His Asp Thr Ser Gly Gln Gln Leu Gly Lys Val
245 250 255

Leu Asp Val Leu Arg Gln Val Asn Gln Ser Lys Val Gln Val Val Val
260 265 270

Leu Phe Ala Ser Ala Arg Ala Val Tyr Ser Leu Phe Ser Tyr Ser Ile
275 280 285

His His Gly Leu Ser Pro Lys Val Trp Val Ala Ser Glu Ser Trp Leu
290 295 300

Thr Ser Asp Leu Val Met Thr Leu Pro Asn Ile Ala Arg Val Gly Thr
305 310 315 320

Val Leu Gly Phe Leu Gln Arg Gly Ala Leu Leu Pro Glu Phe Ser His
325 330 335

Tyr Val Glu Thr His Leu Ala Leu Ala Ala Asp Pro Ala Phe Cys Ala
340 345 350

Ser Leu Asn Ala Glu Leu Asp Leu Glu Glu His Val Met Gly Gln Arg
355 360 365

Cys Pro Arg Cys Asp Asp Ile Met Leu Gln Asn Leu Ser Ser Gly Leu
    370   375   380

Leu Gln Asn Leu Ser Ala Gly Gln Leu His His Gln Ile Phe Ala Thr
385   390   395   400

Tyr Ala Ala Val Tyr Ser Val Ala Gln Ala Leu His Asn Thr Leu Gln
                405   410   415

Cys Asn Val Ser His Cys His Val Ser Glu His Val Leu Pro Trp Gln
            420   425   430

Leu Leu Glu Asn Met Tyr Asn Met Ser Phe His Ala Arg Asp Leu Thr
        435   440   445

Leu Gln Phe Asp Ala Glu Gly Asn Val Asp Met Glu Tyr Asp Leu Lys
    450   455   460

Met Trp Val Trp Gln Ser Pro Thr Pro Val Leu His Thr Val Gly Thr
465   470   475   480

Phe Asn Gly Thr Leu Gln Leu Gln Gln Ser Lys Met Tyr Trp Pro Gly
                485   490   495

Asn Gln Val Pro Val Ser Gln Cys Ser Arg Gln Cys Lys Asp Gly Gln
            500   505   510

Val Arg Arg Val Lys Gly Phe His Ser Cys Cys Tyr Asp Cys Val Asp
        515   520   525

Cys Lys Ala Gly Ser Tyr Arg Lys His Pro Asp Asp Phe Thr Cys Thr
    530   535   540

Pro Cys Asn Gln Asp Gln Trp Ser Pro Glu Lys Ser Thr Ala Cys Leu
545   550   555   560

Pro Arg Arg Pro Lys Phe Leu Ala Trp Gly Glu Pro Val Val Leu Ser
                565   570   575

Leu Leu Leu Leu Leu Cys Leu Val Leu Gly Leu Ala Leu Ala Ala Leu
            580   585   590

Gly Leu Ser Val His His Trp Asp Ser Pro Leu Val Gln Ala Ser Gly
        595   600   605

Gly Ser Gln Phe Cys Phe Gly Leu Ile Cys Leu Gly Leu Phe Cys Leu
    610   615   620

Ser Val Leu Leu Phe Pro Gly Arg Pro Ser Ser Ala Ser Cys Leu Ala
625   630   635   640

Gln Gln Pro Met Ala His Leu Pro Leu Thr Gly Cys Leu Ser Thr Leu
                645   650   655

Phe Leu Gln Ala Ala Glu Thr Phe Val Glu Ser Glu Leu Pro Leu Ser
660 665 670

Trp Ala Asn Trp Leu Cys Ser Tyr Leu Arg Gly Leu Trp Ala Trp Leu
675 680 685

Val Val Leu Leu Ala Thr Phe Val Glu Ala Ala Leu Cys Ala Trp Tyr
690 695 700

Leu Ile Ala Phe Pro Pro Glu Val Val Thr Asp Trp Ser Val Leu Pro
705 710 715 720

Thr Glu Val Leu Glu His Cys His Val Arg Ser Trp Val Ser Leu Gly
725 730 735

Leu Val His Ile Thr Asn Ala Met Leu Ala Phe Leu Cys Phe Leu Gly
740 745 750

Thr Phe Leu Val Gln Ser Gln Pro Gly Arg Tyr Asn Arg Ala Arg Gly
755 760 765

Leu Thr Phe Ala Met Leu Ala Tyr Phe Ile Thr Trp Val Ser Phe Val
770 775 780

Pro Leu Leu Ala Asn Val Gln Val Ala Tyr Gln Pro Ala Val Gln Met
785 790 795 800

Gly Ala Ile Leu Val Cys Ala Leu Gly Ile Leu Val Thr Phe His Leu
805 810 815

Pro Lys Cys Tyr Val Leu Leu Trp Leu Pro Lys Leu Asn Thr Gln Glu
820 825 830

Phe Phe Leu Gly Arg Asn Ala Lys Lys Ala Ala Asp Glu Asn Ser Gly
835 840 845

Gly Gly Glu Ala Ala Gln Gly His Asn Glu
850 855

<210> 14
<211> 858
<212> PRT
<213> Rattus rattus

<400> 14

Met Pro Gly Leu Ala Ile Leu Gly Leu Ser Leu Ala Ala Phe Leu Glu
1 5 10 15

Leu Gly Met Gly Ser Ser Leu Cys Leu Ser Gln Gln Phe Lys Ala Gln
20 25 30

Gly Asp Tyr Ile Leu Gly Gly Leu Phe Pro Leu Gly Thr Thr Glu Glu
35 40 45

Ala Thr Leu Asn Gln Arg Thr Gln Pro Asn Gly Ile Leu Cys Thr Arg
    50   55   60

Phe Ser Pro Leu Gly Leu Phe Leu Ala Met Ala Met Lys Met Ala Val
65   70   75   80

Glu Glu Ile Asn Asn Gly Ser Ala Leu Leu Pro Gly Leu Arg Leu Gly
                85   90   95

Tyr Asp Leu Phe Asp Thr Cys Ser Glu Pro Val Val Thr Met Lys Pro
            100   105   110

Ser Leu Met Phe Met Ala Lys Val Gly Ser Gln Ser Ile Ala Ala Tyr
        115   120   125

Cys Asn Tyr Thr Gln Tyr Gln Pro Arg Val Leu Ala Val Ile Gly Pro
    130   135   140

His Ser Ser Glu Leu Ala Leu Ile Thr Gly Lys Phe Phe Ser Phe Phe
145   150   155   160

Leu Met Pro Gln Val Ser Tyr Ser Ala Ser Met Asp Arg Leu Ser Asp
                165   170   175

Arg Glu Thr Phe Pro Ser Phe Phe Arg Thr Val Pro Ser Asp Arg Val
            180   185   190

Gln Leu Gln Ala Val Val Thr Leu Leu Gln Asn Phe Ser Trp Asn Trp
        195   200   205

Val Ala Ala Leu Gly Ser Asp Asp Asp Tyr Gly Arg Glu Gly Leu Ser
    210   215   220

Ile Phe Ser Gly Leu Ala Asn Ser Arg Gly Ile Cys Ile Ala His Glu
225   230   235   240

Gly Leu Val Pro Gln His Asp Thr Ser Gly Gln Gln Leu Gly Lys Val
                245   250   255

Val Asp Val Leu Arg Gln Val Asn Gln Ser Lys Val Gln Val Val Val
            260   265   270

Leu Phe Ala Ser Ala Arg Ala Val Tyr Ser Leu Phe Ser Tyr Ser Ile
        275   280   285

Leu His Asp Leu Ser Pro Lys Val Trp Val Ala Ser Glu Ser Trp Leu
    290   295   300

Thr Ser Asp Leu Val Met Thr Leu Pro Asn Ile Ala Arg Val Gly Thr
305   310   315   320

Val Leu Gly Phe Leu Gln Arg Gly Ala Leu Leu Pro Glu Phe Ser His
                325   330   335

Tyr Val Glu Thr Arg Leu Ala Leu Ala Ala Asp Pro Thr Phe Cys Ala
340 345 350

Ser Leu Lys Ala Glu Leu Asp Leu Glu Glu Arg Val Met Gly Pro Arg
355 360 365

Cys Ser Gln Cys Asp Tyr Ile Met Leu Gln Asn Leu Ser Ser Gly Leu
370 375 380

Met Gln Asn Leu Ser Ala Gly Gln Leu His His Gln Ile Phe Ala Thr
385 390 395 400

Tyr Ala Ala Val Tyr Ser Val Ala Gln Ala Leu His Asn Thr Leu Gln
405 410 415

Cys Asn Val Ser His Cys His Thr Ser Glu Pro Val Gln Pro Trp Gln
420 425 430

Leu Leu Glu Asn Met Tyr Asn Met Ser Phe Arg Ala Arg Asp Leu Thr
435 440 445

Leu Gln Phe Asp Ala Lys Gly Ser Val Asp Met Glu Tyr Asp Leu Lys
450 455 460

Met Trp Val Trp Gln Ser Pro Thr Pro Val Leu His Thr Val Gly Thr
465 470 475 480

Phe Asn Gly Thr Leu Gln Leu Gln His Ser Lys Met Tyr Trp Pro Gly
485 490 495

Asn Gln Val Pro Val Ser Gln Cys Ser Arg Gln Cys Lys Asp Gly Gln
500 505 510

Val Arg Arg Val Lys Gly Phe His Ser Cys Cys Tyr Asp Cys Val Asp
515 520 525

Cys Lys Ala Gly Ser Tyr Arg Lys His Pro Asp Asp Phe Thr Cys Thr
530 535 540

Pro Cys Gly Lys Asp Gln Trp Ser Pro Glu Lys Ser Thr Thr Cys Leu
545 550 555 560

Pro Arg Arg Pro Lys Phe Leu Ala Trp Gly Glu Pro Ala Val Leu Ser
565 570 575

Leu Leu Leu Leu Leu Cys Leu Val Leu Gly Leu Thr Leu Ala Ala Leu
580 585 590

Gly Leu Phe Val His Tyr Trp Asp Ser Pro Leu Val Gln Ala Ser Gly
595 600 605

Gly Ser Leu Phe Cys Phe Gly Leu Ile Cys Leu Gly Leu Phe Cys Leu

610 615 620

Ser Val Leu Leu Phe Pro Gly Arg Pro Arg Ser Ala Ser Cys Leu Ala
625 630 635 640

Gln Gln Pro Met Ala His Leu Pro Leu Thr Gly Cys Leu Ser Thr Leu
645 650 655

Phe Leu Gln Ala Ala Glu Ile Phe Val Glu Ser Glu Leu Pro Leu Ser
660 665 670

Trp Ala Asn Trp Leu Cys Ser Tyr Leu Arg Gly Pro Trp Ala Trp Leu
675 680 685

Val Val Leu Leu Ala Thr Leu Val Glu Ala Ala Leu Cys Ala Trp Tyr
690 695 700

Leu Met Ala Phe Pro Pro Glu Val Val Thr Asp Trp Gln Val Leu Pro
705 710 715 720

Thr Glu Val Leu Glu His Cys Arg Met Arg Ser Trp Val Ser Leu Gly
725 730 735

Leu Val His Ile Thr Asn Ala Val Leu Ala Phe Leu Cys Phe Leu Gly
740 745 750

Thr Phe Leu Val Gln Ser Gln Pro Gly Arg Tyr Asn Arg Ala Arg Gly
755 760 765

Leu Thr Phe Ala Met Leu Ala Tyr Phe Ile Ile Trp Val Ser Phe Val
770 775 780

Pro Leu Leu Ala Asn Val Gln Val Ala Tyr Gln Pro Ala Val Gln Met
785 790 795 800

Gly Ala Ile Leu Phe Cys Ala Leu Gly Ile Leu Ala Thr Phe His Leu
805 810 815

Pro Lys Cys Tyr Val Leu Leu Trp Leu Pro Glu Leu Asn Thr Gln Glu
820 825 830

Phe Phe Leu Gly Arg Ser Pro Lys Glu Ala Ser Asp Gly Asn Ser Gly
835 840 845

Ser Ser Glu Ala Thr Arg Gly His Ser Glu
850 855

<210> 15
<211> 842
<212> PRT
<213> Mus musculus

<400> 15

Met Leu Phe Trp Ala Ala His Leu Leu Leu Ser Leu Gln Leu Ala Val

1 5 10 15

Ala Tyr Cys Trp Ala Phe Ser Cys Gln Arg Thr Glu Ser Ser Pro Gly
20 25 30

Phe Ser Leu Pro Gly Asp Phe Leu Leu Ala Gly Leu Phe Ser Leu His
35 40 45

Ala Asp Cys Leu Gln Val Arg His Arg Pro Leu Val Thr Ser Cys Asp
50 55 60

Arg Ser Asp Ser Phe Asn Gly His Gly Tyr His Leu Phe Gln Ala Met
65 70 75 80

Arg Phe Thr Val Glu Glu Ile Asn Asn Ser Thr Ala Leu Leu Pro Asn
85 90 95

Ile Thr Leu Gly Tyr Glu Leu Tyr Asp Val Cys Ser Glu Ser Ser Asn
100 105 110

Val Tyr Ala Thr Leu Arg Val Leu Ala Gln Gln Gly Thr Gly His Leu
115 120 125

Glu Met Gln Arg Asp Leu Arg Asn His Ser Ser Lys Val Val Ala Leu
130 135 140

Ile Gly Pro Asp Asn Thr Asp His Ala Val Thr Thr Ala Ala Leu Leu
145 150 155 160

Ser Pro Phe Leu Met Pro Leu Val Ser Tyr Glu Ala Ser Ser Val Ile
165 170 175

Leu Ser Gly Lys Arg Lys Phe Pro Ser Phe Leu Arg Thr Ile Pro Ser
180 185 190

Asp Lys Tyr Gln Val Glu Val Ile Val Arg Leu Leu Gln Ser Phe Gly
195 200 205

Trp Val Trp Ile Ser Leu Val Gly Ser Tyr Gly Asp Tyr Gly Gln Leu
210 215 220

Gly Val Gln Ala Leu Glu Glu Leu Ala Thr Pro Arg Gly Ile Cys Val
225 230 235 240

Ala Phe Lys Asp Val Val Pro Leu Ser Ala Gln Ala Gly Asp Pro Arg
245 250 255

Met Gln Arg Met Met Leu Arg Leu Ala Arg Ala Arg Thr Thr Val Val
260 265 270

Val Val Phe Ser Asn Arg His Leu Ala Gly Val Phe Phe Arg Ser Val
275 280 285

```
Val Leu Ala Asn Leu Thr Gly Lys Val Trp Ile Ala Ser Glu Asp Trp
    290                 295                 300

Ala Ile Ser Thr Tyr Ile Thr Asn Val Pro Gly Ile Gln Gly Ile Gly
305                 310                 315                 320

Thr Val Leu Gly Val Ala Ile Gln Gln Arg Gln Val Pro Gly Leu Lys
                325                 330                 335

Glu Phe Glu Glu Ser Tyr Val Gln Ala Val Met Gly Ala Pro Arg Thr
            340                 345                 350

Cys Pro Glu Gly Ser Trp Cys Gly Thr Asn Gln Leu Cys Arg Glu Cys
        355                 360                 365

His Ala Phe Thr Thr Trp Asn Met Pro Glu Leu Gly Ala Phe Ser Met
    370                 375                 380

Ser Ala Ala Tyr Asn Val Tyr Glu Ala Val Tyr Ala Val Ala His Gly
385                 390                 395                 400

Leu His Gln Leu Leu Gly Cys Thr Ser Gly Thr Cys Ala Arg Gly Pro
                405                 410                 415

Val Tyr Pro Trp Gln Leu Leu Gln Gln Ile Tyr Lys Val Asn Phe Leu
            420                 425                 430

Leu His Lys Lys Thr Val Ala Phe Asp Asp Lys Gly Asp Pro Leu Gly
        435                 440                 445

Tyr Tyr Asp Ile Ile Ala Trp Asp Trp Asn Gly Pro Glu Trp Thr Phe
    450                 455                 460

Glu Val Ile Gly Ser Ala Ser Leu Ser Pro Val His Leu Asp Ile Asn
465                 470                 475                 480

Lys Thr Lys Ile Gln Trp His Gly Lys Asn Asn Gln Val Pro Val Ser
                485                 490                 495

Val Cys Thr Arg Asp Cys Leu Glu Gly His His Arg Leu Val Met Gly
            500                 505                 510

Ser His His Cys Cys Phe Glu Cys Met Pro Cys Glu Ala Gly Thr Phe
        515                 520                 525

Leu Asn Thr Ser Glu Leu His Thr Cys Gln Pro Cys Gly Thr Glu Glu
    530                 535                 540

Trp Ala Pro Glu Gly Ser Ser Ala Cys Phe Ser Arg Thr Val Glu Phe
545                 550                 555                 560

Leu Gly Trp His Glu Pro Ile Ser Leu Val Leu Leu Ala Ala Asn Thr
                565                 570                 575
```

Leu Leu Leu Leu Leu Leu Ile Gly Thr Ala Gly Leu Phe Ala Trp Arg
580 585 590

Leu His Thr Pro Val Val Arg Ser Ala Gly Gly Arg Leu Cys Phe Leu
595 600 605

Met Leu Gly Ser Leu Val Ala Gly Ser Cys Ser Leu Tyr Ser Phe Phe
610 615 620

Gly Lys Pro Thr Val Pro Ala Cys Leu Leu Arg Gln Pro Leu Phe Ser
625 630 635 640

Leu Gly Phe Ala Ile Phe Leu Ser Cys Leu Thr Ile Arg Ser Phe Gln
645 650 655

Leu Val Ile Ile Phe Lys Phe Ser Thr Lys Val Pro Thr Phe Tyr His
660 665 670

Thr Trp Ala Gln Asn His Gly Ala Gly Ile Phe Val Ile Val Ser Ser
675 680 685

Thr Val His Leu Phe Leu Cys Leu Thr Trp Leu Ala Met Trp Thr Pro
690 695 700

Arg Pro Thr Arg Glu Tyr Gln Arg Phe Pro His Leu Val Ile Leu Glu
705 710 715 720

Cys Thr Glu Val Asn Ser Val Gly Phe Leu Val Ala Phe Ala His Asn
725 730 735

Ile Leu Leu Ser Ile Ser Thr Phe Val Cys Ser Tyr Leu Gly Lys Glu
740 745 750

Leu Pro Glu Asn Tyr Asn Glu Ala Lys Cys Val Thr Phe Ser Leu Leu
755 760 765

Leu His Phe Val Ser Trp Ile Ala Phe Phe Thr Met Ser Ser Ile Tyr
770 775 780

Gln Gly Ser Tyr Leu Pro Ala Val Asn Val Leu Ala Gly Leu Ala Thr
785 790 795 800

Leu Ser Gly Gly Phe Ser Gly Tyr Phe Leu Pro Lys Cys Tyr Val Ile
805 810 815

Leu Cys Arg Pro Glu Leu Asn Asn Thr Glu His Phe Gln Ala Ser Ile
820 825 830

Gln Asp Tyr Thr Arg Arg Cys Gly Thr Thr
835 840

<210> 16
<211> 840

<212> PRT
<213> Rattus rattus

<400> 16

Met Leu Phe Trp Ala Ala His Leu Leu Leu Ser Leu Gln Leu Val Tyr
1 5 10 15

Cys Trp Ala Phe Ser Cys Gln Arg Thr Glu Ser Ser Pro Gly Phe Ser
20 25 30

Leu Pro Gly Asp Phe Leu Leu Ala Gly Leu Phe Ser Leu His Gly Asp
35 40 45

Cys Leu Gln Val Arg His Arg Pro Leu Val Thr Ser Cys Asp Arg Pro
50 55 60

Asp Ser Phe Asn Gly His Gly Tyr His Leu Phe Gln Ala Met Arg Phe
65 70 75 80

Thr Val Glu Glu Ile Asn Asn Ser Ser Ala Leu Leu Pro Asn Ile Thr
85 90 95

Leu Gly Tyr Glu Leu Tyr Asp Val Cys Ser Glu Ser Ala Asn Val Tyr
100 105 110

Ala Thr Leu Arg Val Leu Ala Leu Gln Gly Pro Arg His Ile Glu Ile
115 120 125

Gln Lys Asp Leu Arg Asn His Ser Ser Lys Val Val Ala Phe Ile Gly
130 135 140

Pro Asp Asn Thr Asp His Ala Val Thr Thr Ala Ala Leu Leu Gly Pro
145 150 155 160

Phe Leu Met Pro Leu Val Ser Tyr Glu Ala Ser Ser Val Val Leu Ser
165 170 175

Ala Lys Arg Lys Phe Pro Ser Phe Leu Arg Thr Val Pro Ser Asp Arg
180 185 190

His Gln Val Glu Val Met Val Gln Leu Leu Gln Ser Phe Gly Trp Val
195 200 205

Trp Ile Ser Leu Ile Gly Ser Tyr Gly Asp Tyr Gly Gln Leu Gly Val
210 215 220

Gln Ala Leu Glu Glu Leu Ala Val Pro Arg Gly Ile Cys Val Ala Phe
225 230 235 240

Lys Asp Ile Val Pro Phe Ser Ala Arg Val Gly Asp Pro Arg Met Gln
245 250 255

Ser Met Met Gln His Leu Ala Gln Ala Arg Thr Thr Val Val Val Val
260 265 270

```
Phe Ser Asn Arg His Leu Ala Arg Val Phe Phe Arg Ser Val Val Leu
        275                 280                 285

Ala Asn Leu Thr Gly Lys Val Trp Val Ala Ser Glu Asp Trp Ala Ile
    290                 295                 300

Ser Thr Tyr Ile Thr Ser Val Thr Gly Ile Gln Gly Ile Gly Thr Val
305                 310                 315                 320

Leu Gly Val Ala Val Gln Gln Arg Gln Val Pro Gly Leu Lys Glu Phe
                325                 330                 335

Glu Glu Ser Tyr Val Arg Ala Val Thr Ala Ala Pro Ser Ala Cys Pro
            340                 345                 350

Glu Gly Ser Trp Cys Ser Thr Asn Gln Leu Cys Arg Glu Cys His Thr
        355                 360                 365

Phe Thr Thr Arg Asn Met Pro Thr Leu Gly Ala Phe Ser Met Ser Ala
    370                 375                 380

Ala Tyr Arg Val Tyr Glu Ala Val Tyr Ala Val Ala His Gly Leu His
385                 390                 395                 400

Gln Leu Leu Gly Cys Thr Ser Glu Ile Cys Ser Arg Gly Pro Val Tyr
                405                 410                 415

Pro Trp Gln Leu Leu Gln Gln Ile Tyr Lys Val Asn Phe Leu Leu His
            420                 425                 430

Glu Asn Thr Val Ala Phe Asp Asp Asn Gly Asp Thr Leu Gly Tyr Tyr
        435                 440                 445

Asp Ile Ile Ala Trp Asp Trp Asn Gly Pro Glu Trp Thr Phe Glu Ile
    450                 455                 460

Ile Gly Ser Ala Ser Leu Ser Pro Val His Leu Asp Ile Asn Lys Thr
465                 470                 475                 480

Lys Ile Gln Trp His Gly Lys Asn Asn Gln Val Pro Val Ser Val Cys
                485                 490                 495

Thr Thr Asp Cys Leu Ala Gly His His Arg Val Val Val Gly Ser His
            500                 505                 510

His Cys Cys Phe Glu Cys Val Pro Cys Glu Ala Gly Thr Phe Leu Asn
        515                 520                 525

Met Ser Glu Leu His Ile Cys Gln Pro Cys Gly Thr Glu Glu Trp Ala
    530                 535                 540

Pro Lys Glu Ser Thr Thr Cys Phe Pro Arg Thr Val Glu Phe Leu Ala
```

545 550 555 560

Trp His Glu Pro Ile Ser Leu Val Leu Ile Ala Ala Asn Thr Leu Leu
565 570 575

Leu Leu Leu Leu Val Gly Thr Ala Gly Leu Phe Ala Trp His Phe His
580 585 590

Thr Pro Val Val Arg Ser Ala Gly Gly Arg Leu Cys Phe Leu Met Leu
595 600 605

Gly Ser Leu Val Ala Gly Ser Cys Ser Phe Tyr Ser Phe Phe Gly Glu
610 615 620

Pro Thr Val Pro Ala Cys Leu Leu Arg Gln Pro Leu Phe Ser Leu Gly
625 630 635 640

Phe Ala Ile Phe Leu Ser Cys Leu Thr Ile Arg Ser Phe Gln Leu Val
645 650 655

Ile Ile Phe Lys Phe Ser Thr Lys Val Pro Thr Phe Tyr Arg Thr Trp
660 665 670

Ala Gln Asn His Gly Ala Gly Leu Phe Val Ile Val Ser Ser Thr Val
675 680 685

His Leu Leu Ile Cys Leu Thr Trp Leu Val Met Trp Thr Pro Arg Pro
690 695 700

Thr Arg Glu Tyr Gln Arg Phe Pro His Leu Val Ile Leu Glu Cys Thr
705 710 715 720

Glu Val Asn Ser Val Gly Phe Leu Leu Ala Phe Thr His Asn Ile Leu
725 730 735

Leu Ser Ile Ser Thr Phe Val Cys Ser Tyr Leu Gly Lys Glu Leu Pro
740 745 750

Glu Asn Tyr Asn Glu Ala Lys Cys Val Thr Phe Ser Leu Leu Leu Asn
755 760 765

Phe Val Ser Trp Ile Ala Phe Phe Thr Met Ala Ser Ile Tyr Gln Gly
770 775 780

Ser Tyr Leu Pro Ala Val Asn Val Leu Ala Gly Leu Thr Thr Leu Ser
785 790 795 800

Gly Gly Phe Ser Gly Tyr Phe Leu Pro Lys Cys Tyr Val Ile Leu Cys
805 810 815

Arg Pro Glu Leu Asn Asn Thr Glu His Phe Gln Ala Ser Ile Gln Asp
820 825 830

Tyr Thr Arg Arg Cys Gly Thr Thr
835 840

<210> 17
<211> 841
<212> PRT
<213> Homo sapiens

<400> 17

Met Leu Leu Cys Thr Ala Arg Leu Val Gly Leu Gln Leu Leu Ile Ser
1 5 10 15

Cys Cys Trp Ala Phe Ala Cys His Ser Thr Glu Ser Ser Pro Asp Phe
20 25 30

Thr Leu Pro Gly Asp Tyr Leu Leu Ala Gly Leu Phe Pro Leu His Ser
35 40 45

Gly Cys Leu Gln Val Arg His Arg Pro Glu Val Thr Leu Cys Asp Arg
50 55 60

Ser Cys Ser Phe Asn Glu His Gly Tyr His Leu Phe Gln Ala Met Arg
65 70 75 80

Leu Gly Val Glu Glu Ile Asn Asn Ser Thr Ala Leu Leu Pro Asn Ile
85 90 95

Thr Leu Gly Tyr Gln Leu Tyr Asp Val Cys Ser Asp Ser Ala Asn Val
100 105 110

Tyr Ala Thr Leu Arg Val Leu Ser Leu Pro Gly Gln His His Ile Glu
115 120 125

Leu Gln Gly Asp Leu Leu His Tyr Ser Pro Thr Val Leu Ala Val Ile
130 135 140

Gly Pro Asp Ser Thr Asn Arg Ala Ala Thr Thr Ala Ala Leu Leu Ser
145 150 155 160

Pro Phe Leu Val Pro Met Ile Ser Tyr Ala Ala Ser Ser Glu Thr Leu
165 170 175

Ser Val Lys Arg Gln Tyr Pro Ser Phe Leu Arg Thr Ile Pro Asn Asp
180 185 190

Lys Tyr Gln Val Glu Thr Met Val Leu Leu Leu Gln Lys Phe Gly Trp
195 200 205

Thr Trp Ile Ser Leu Val Gly Ser Ser Asp Asp Tyr Gly Gln Leu Gly
210 215 220

Val Gln Ala Leu Glu Asn Gln Ala Thr Gly Gln Gly Ile Cys Ile Ala
225 230 235 240

Phe Lys Asp Ile Met Pro Phe Ser Ala Gln Val Gly Asp Glu Arg Met
                245   250   255

Gln Cys Leu Met Arg His Leu Ala Gln Ala Gly Ala Thr Val Val Val
            260   265   270

Val Phe Ser Ser Arg Gln Leu Ala Arg Val Phe Phe Glu Ser Val Val
        275   280   285

Leu Thr Asn Leu Thr Gly Lys Val Trp Val Ala Ser Glu Ala Trp Ala
    290   295   300

Leu Ser Arg His Ile Thr Gly Val Pro Gly Ile Gln Arg Ile Gly Met
305   310   315   320

Val Leu Gly Val Ala Ile Gln Lys Arg Ala Val Pro Gly Leu Lys Ala
                325   330   335

Phe Glu Glu Ala Tyr Ala Arg Ala Asp Lys Lys Ala Pro Arg Pro Cys
            340   345   350

His Lys Gly Ser Trp Cys Ser Ser Asn Gln Leu Cys Arg Glu Cys Gln
        355   360   365

Ala Phe Met Ala His Thr Met Pro Lys Leu Lys Ala Phe Ser Met Ser
    370   375   380

Ser Ala Tyr Asn Ala Tyr Arg Ala Val Tyr Ala Val Ala His Gly Leu
385   390   395   400

His Gln Leu Leu Gly Cys Ala Ser Gly Ala Cys Ser Arg Gly Arg Val
                405   410   415

Tyr Pro Trp Gln Leu Leu Glu Gln Ile His Lys Val His Phe Leu Leu
            420   425   430

His Lys Asp Thr Val Ala Phe Asn Asp Asn Arg Asp Pro Leu Ser Ser
        435   440   445

Tyr Asn Ile Ile Ala Trp Asp Trp Asn Gly Pro Lys Trp Thr Phe Thr
    450   455   460

Val Leu Gly Ser Ser Thr Trp Ser Pro Val Gln Leu Asn Ile Asn Glu
465   470   475   480

Thr Lys Ile Gln Trp His Gly Lys Asp Asn Gln Val Pro Lys Ser Val
                485   490   495

Cys Ser Ser Asp Cys Leu Glu Gly His Gln Arg Val Val Thr Gly Phe
            500   505   510

His His Cys Cys Phe Glu Cys Val Pro Cys Gly Ala Gly Thr Phe Leu
        515   520   525

```
Asn Lys Ser Asp Leu Tyr Arg Cys Gln Pro Cys Gly Lys Glu Glu Trp
    530                 535                 540

Ala Pro Glu Gly Ser Gln Thr Cys Phe Pro Arg Thr Val Val Phe Leu
545                 550                 555                 560

Ala Leu Arg Glu His Thr Ser Trp Val Leu Leu Ala Ala Asn Thr Leu
                565                 570                 575

Leu Leu Leu Leu Leu Leu Gly Thr Ala Gly Leu Phe Ala Trp His Leu
            580                 585                 590

Asp Thr Pro Val Val Arg Ser Ala Gly Gly Arg Leu Cys Phe Leu Met
        595                 600                 605

Leu Gly Ser Leu Ala Ala Gly Ser Gly Ser Leu Tyr Gly Phe Phe Gly
    610                 615                 620

Glu Pro Thr Arg Pro Ala Cys Leu Leu Arg Gln Ala Leu Phe Ala Leu
625                 630                 635                 640

Gly Phe Thr Ile Phe Leu Ser Cys Leu Thr Val Arg Ser Phe Gln Leu
                645                 650                 655

Ile Ile Ile Phe Lys Phe Ser Thr Lys Val Pro Thr Phe Tyr His Ala
            660                 665                 670

Trp Val Gln Asn His Gly Ala Gly Leu Phe Val Met Ile Ser Ser Ala
        675                 680                 685

Ala Gln Leu Leu Ile Cys Leu Thr Trp Leu Val Val Trp Thr Pro Leu
    690                 695                 700

Pro Ala Arg Glu Tyr Gln Arg Phe Pro His Leu Val Met Leu Glu Cys
705                 710                 715                 720

Thr Glu Thr Asn Ser Leu Gly Phe Ile Leu Ala Phe Leu Tyr Asn Gly
                725                 730                 735

Leu Leu Ser Ile Ser Ala Phe Ala Cys Ser Tyr Leu Gly Lys Asp Leu
            740                 745                 750

Pro Glu Asn Tyr Asn Glu Ala Lys Cys Val Thr Phe Ser Leu Leu Phe
        755                 760                 765

Asn Phe Val Ser Trp Ile Ala Phe Phe Thr Thr Ala Ser Val Tyr Asp
    770                 775                 780

Gly Lys Tyr Leu Pro Ala Ala Asn Met Met Ala Gly Leu Ser Ser Leu
785                 790                 795                 800

Ser Ser Gly Phe Gly Gly Tyr Phe Leu Pro Lys Cys Tyr Val Ile Leu
                805                 810                 815
```

Cys Arg Pro Asp Leu Asn Ser Thr Glu His Phe Gln Ala Ser Ile Gln
820 825 830

Asp Tyr Thr Arg Arg Cys Gly Ser Thr
835 840

<210> 18
<211> 843
<212> PRT
<213> Mus musculus

<400> 18

Met Gly Pro Gln Ala Arg Thr Leu His Leu Leu Phe Leu Leu Leu His
1 5 10 15

Ala Leu Pro Lys Pro Val Met Leu Val Gly Asn Ser Asp Phe His Leu
20 25 30

Ala Gly Asp Tyr Leu Leu Gly Gly Leu Phe Thr Leu His Ala Asn Val
35 40 45

Lys Ser Val Ser His Leu Ser Tyr Leu Gln Val Pro Lys Cys Asn Glu
50 55 60

Tyr Asn Met Lys Val Leu Gly Tyr Asn Leu Met Gln Ala Met Arg Phe
65 70 75 80

Ala Val Glu Glu Ile Asn Asn Cys Ser Ser Leu Leu Pro Gly Val Leu
85 90 95

Leu Gly Tyr Glu Met Val Asp Val Cys Tyr Leu Ser Asn Asn Ile Gln
100 105 110

Pro Gly Leu Tyr Phe Leu Ser Gln Ile Asp Asp Phe Leu Pro Ile Leu
115 120 125

Lys Asp Tyr Ser Gln Tyr Arg Pro Gln Val Val Ala Val Ile Gly Pro
130 135 140

Asp Asn Ser Glu Ser Ala Ile Thr Val Ser Asn Ile Leu Ser Tyr Phe
145 150 155 160

Leu Val Pro Gln Val Thr Tyr Ser Ala Ile Thr Asp Lys Leu Arg Asp
165 170 175

Lys Arg Arg Phe Pro Ala Met Leu Arg Thr Val Pro Ser Ala Thr His
180 185 190

His Ile Glu Ala Met Val Gln Leu Met Val His Phe Gln Trp Asn Trp
195 200 205

Ile Val Val Leu Val Ser Asp Asp Asp Tyr Gly Arg Glu Asn Ser His
210 215 220

Leu Leu Ser Gln Arg Leu Thr Asn Thr Gly Asp Ile Cys Ile Ala Phe
225   230   235   240

Gln Glu Val Leu Pro Val Pro Glu Pro Asn Gln Ala Val Arg Pro Glu
                245   250   255

Glu Gln Asp Gln Leu Asp Asn Ile Leu Asp Lys Leu Arg Arg Thr Ser
            260   265   270

Ala Arg Val Val Val Ile Phe Ser Pro Glu Leu Ser Leu His Asn Phe
        275   280   285

Phe Arg Glu Val Leu Arg Trp Asn Phe Thr Gly Phe Val Trp Ile Ala
    290   295   300

Ser Glu Ser Trp Ala Ile Asp Pro Val Leu His Asn Leu Thr Glu Leu
305   310   315   320

Arg His Thr Gly Thr Phe Leu Gly Val Thr Ile Gln Arg Val Ser Ile
                325   330   335

Pro Gly Phe Ser Gln Phe Arg Val Arg His Asp Lys Pro Glu Tyr Pro
            340   345   350

Met Pro Asn Glu Thr Ser Leu Arg Thr Thr Cys Asn Gln Asp Cys Asp
        355   360   365

Ala Cys Met Asn Ile Thr Glu Ser Phe Asn Asn Val Leu Met Leu Ser
    370   375   380

Gly Glu Arg Val Val Tyr Ser Val Tyr Ser Ala Val Tyr Ala Val Ala
385   390   395   400

His Thr Leu His Arg Leu Leu His Cys Asn Gln Val Arg Cys Thr Lys
                405   410   415

Gln Ile Val Tyr Pro Trp Gln Leu Leu Arg Glu Ile Trp His Val Asn
            420   425   430

Phe Thr Leu Leu Gly Asn Gln Leu Phe Phe Asp Glu Gln Gly Asp Met
        435   440   445

Pro Met Leu Leu Asp Ile Ile Gln Trp Gln Trp Gly Leu Ser Gln Asn
    450   455   460

Pro Phe Gln Ser Ile Ala Ser Tyr Ser Pro Thr Glu Thr Arg Leu Thr
465   470   475   480

Tyr Ile Ser Asn Val Ser Trp Tyr Thr Pro Asn Asn Thr Val Pro Ile
                485   490   495

Ser Met Cys Ser Lys Ser Cys Gln Pro Gly Gln Met Lys Lys Pro Ile

500 505 510

Gly Leu His Pro Cys Cys Phe Glu Cys Val Asp Cys Pro Pro Gly Thr
515 520 525

Tyr Leu Asn Arg Ser Val Asp Glu Phe Asn Cys Leu Ser Cys Pro Gly
530 535 540

Ser Met Trp Ser Tyr Lys Asn Asn Ile Ala Cys Phe Lys Arg Arg Leu
545 550 555 560

Ala Phe Leu Glu Trp His Glu Val Pro Thr Ile Val Val Thr Ile Leu
565 570 575

Ala Ala Leu Gly Phe Ile Ser Thr Leu Ala Ile Leu Leu Ile Phe Trp
580 585 590

Arg His Phe Gln Thr Pro Met Val Arg Ser Ala Gly Gly Pro Met Cys
595 600 605

Phe Leu Met Leu Val Pro Leu Leu Leu Ala Phe Gly Met Val Pro Val
610 615 620

Tyr Val Gly Pro Pro Thr Val Phe Ser Cys Phe Cys Arg Gln Ala Phe
625 630 635 640

Phe Thr Val Cys Phe Ser Val Cys Leu Ser Cys Ile Thr Val Arg Ser
645 650 655

Phe Gln Ile Val Cys Val Phe Lys Met Ala Arg Arg Leu Pro Ser Ala
660 665 670

Tyr Gly Phe Trp Met Arg Tyr His Gly Pro Tyr Val Phe Val Ala Phe
675 680 685

Ile Thr Ala Val Lys Val Ala Leu Val Ala Gly Asn Met Leu Ala Thr
690 695 700

Thr Ile Asn Pro Ile Gly Arg Thr Asp Pro Asp Asp Pro Asn Ile Ile
705 710 715 720

Ile Leu Ser Cys His Pro Asn Tyr Arg Asn Gly Leu Leu Phe Asn Thr
725 730 735

Ser Met Asp Leu Leu Leu Ser Val Leu Gly Phe Ser Phe Ala Tyr Val
740 745 750

Gly Lys Glu Leu Pro Thr Asn Tyr Asn Glu Ala Lys Phe Ile Thr Leu
755 760 765

Ser Met Thr Phe Ser Phe Thr Ser Ser Ile Ser Leu Cys Thr Phe Met
770 775 780

Ser Val His Asp Gly Val Leu Val Thr Ile Met Asp Leu Leu Val Thr
785 790 795 800

Val Leu Asn Phe Leu Ala Ile Gly Leu Gly Tyr Phe Gly Pro Lys Cys
805 810 815

Tyr Met Ile Leu Phe Tyr Pro Glu Arg Asn Thr Ser Ala Tyr Phe Asn
820 825 830

Ser Met Ile Gln Gly Tyr Thr Met Arg Lys Ser
835 840

<210> 19
<211> 843
<212> PRT
<213> Rattus rattus

<400> 19

Met Gly Pro Gln Ala Arg Thr Leu Cys Leu Leu Ser Leu Leu Leu His
1 5 10 15

Val Leu Pro Lys Pro Gly Lys Leu Val Glu Asn Ser Asp Phe His Leu
20 25 30

Ala Gly Asp Tyr Leu Leu Gly Gly Leu Phe Thr Leu His Ala Asn Val
35 40 45

Lys Ser Ile Ser His Leu Ser Tyr Leu Gln Val Pro Lys Cys Asn Glu
50 55 60

Phe Thr Met Lys Val Leu Gly Tyr Asn Leu Met Gln Ala Met Arg Phe
65 70 75 80

Ala Val Glu Glu Ile Asn Asn Cys Ser Ser Leu Leu Pro Gly Val Leu
85 90 95

Leu Gly Tyr Glu Met Val Asp Val Cys Tyr Leu Ser Asn Asn Ile His
100 105 110

Pro Gly Leu Tyr Phe Leu Ala Gln Asp Asp Asp Leu Leu Pro Ile Leu
115 120 125

Lys Asp Tyr Ser Gln Tyr Met Pro His Val Val Ala Val Ile Gly Pro
130 135 140

Asp Asn Ser Glu Ser Ala Ile Thr Val Ser Asn Ile Leu Ser His Phe
145 150 155 160

Leu Ile Pro Gln Ile Thr Tyr Ser Ala Ile Ser Asp Lys Leu Arg Asp
165 170 175

Lys Arg His Phe Pro Ser Met Leu Arg Thr Val Pro Ser Ala Thr His
180 185 190

His Ile Glu Ala Met Val Gln Leu Met Val His Phe Gln Trp Asn Trp
195 200 205

Ile Val Val Leu Val Ser Asp Asp Asp Tyr Gly Arg Glu Asn Ser His
210 215 220

Leu Leu Ser Gln Arg Leu Thr Lys Thr Ser Asp Ile Cys Ile Ala Phe
225 230 235 240

Gln Glu Val Leu Pro Ile Pro Glu Ser Ser Gln Val Met Arg Ser Glu
245 250 255

Glu Gln Arg Gln Leu Asp Asn Ile Leu Asp Lys Leu Arg Arg Thr Ser
260 265 270

Ala Arg Val Val Val Val Phe Ser Pro Glu Leu Ser Leu Tyr Ser Phe
275 280 285

Phe His Glu Val Leu Arg Trp Asn Phe Thr Gly Phe Val Trp Ile Ala
290 295 300

Ser Glu Ser Trp Ala Ile Asp Pro Val Leu His Asn Leu Thr Glu Leu
305 310 315 320

Arg His Thr Gly Thr Phe Leu Gly Val Thr Ile Gln Arg Val Ser Ile
325 330 335

Pro Gly Phe Ser Gln Phe Arg Val Arg Arg Asp Lys Pro Gly Tyr Pro
340 345 350

Val Pro Asn Thr Thr Asn Leu Arg Thr Thr Cys Asn Gln Asp Cys Asp
355 360 365

Ala Cys Leu Asn Thr Thr Lys Ser Phe Asn Asn Ile Leu Ile Leu Ser
370 375 380

Gly Glu Arg Val Val Tyr Ser Val Tyr Ser Ala Val Tyr Ala Val Ala
385 390 395 400

His Ala Leu His Arg Leu Leu Gly Cys Asn Arg Val Arg Cys Thr Lys
405 410 415

Gln Lys Val Tyr Pro Trp Gln Leu Leu Arg Glu Ile Trp His Val Asn
420 425 430

Phe Thr Leu Leu Gly Asn Arg Leu Phe Phe Asp Gln Gln Gly Asp Met
435 440 445

Pro Met Leu Leu Asp Ile Ile Gln Trp Gln Trp Asp Leu Ser Gln Asn
450 455 460

Pro Phe Gln Ser Ile Ala Ser Tyr Ser Pro Thr Ser Lys Arg Leu Thr
465 470 475 480

Tyr Ile Asn Asn Val Ser Trp Tyr Thr Pro Asn Asn Thr Val Pro Val
                485   490   495

Ser Met Cys Ser Lys Ser Cys Gln Pro Gly Gln Met Lys Lys Ser Val
            500   505   510

Gly Leu His Pro Cys Cys Phe Glu Cys Leu Asp Cys Met Pro Gly Thr
        515   520   525

Tyr Leu Asn Arg Ser Ala Asp Glu Phe Asn Cys Leu Ser Cys Pro Gly
    530   535   540

Ser Met Trp Ser Tyr Lys Asn Asp Ile Thr Cys Phe Gln Arg Arg Pro
545   550   555   560

Thr Phe Leu Glu Trp His Glu Val Pro Thr Ile Val Val Ala Ile Leu
                565   570   575

Ala Ala Leu Gly Phe Phe Ser Thr Leu Ala Ile Leu Phe Ile Phe Trp
            580   585   590

Arg His Phe Gln Thr Pro Met Val Arg Ser Ala Gly Gly Pro Met Cys
        595   600   605

Phe Leu Met Leu Val Pro Leu Leu Leu Ala Phe Gly Met Val Pro Val
    610   615   620

Tyr Val Gly Pro Pro Thr Val Phe Ser Cys Phe Cys Arg Gln Ala Phe
625   630   635   640

Phe Thr Val Cys Phe Ser Ile Cys Leu Ser Cys Ile Thr Val Arg Ser
                645   650   655

Phe Gln Ile Val Cys Val Phe Lys Met Ala Arg Arg Leu Pro Ser Ala
            660   665   670

Tyr Ser Phe Trp Met Arg Tyr His Gly Pro Tyr Val Phe Val Ala Phe
        675   680   685

Ile Thr Ala Ile Lys Val Ala Leu Val Val Gly Asn Met Leu Ala Thr
    690   695   700

Thr Ile Asn Pro Ile Gly Arg Thr Asp Pro Asp Asp Pro Asn Ile Met
705   710   715   720

Ile Leu Ser Cys His Pro Asn Tyr Arg Asn Gly Leu Leu Phe Asn Thr
                725   730   735

Ser Met Asp Leu Leu Leu Ser Val Leu Gly Phe Ser Phe Ala Tyr Met
            740   745   750

Gly Lys Glu Leu Pro Thr Asn Tyr Asn Glu Ala Lys Phe Ile Thr Leu
        755   760   765

Ser Met Thr Phe Ser Phe Thr Ser Ser Ile Ser Leu Cys Thr Phe Met
770 775 780

Ser Val His Asp Gly Val Leu Val Thr Ile Met Asp Leu Leu Val Thr
785 790 795 800

Val Leu Asn Phe Leu Ala Ile Gly Leu Gly Tyr Phe Gly Pro Lys Cys
805 810 815

Tyr Met Ile Leu Phe Tyr Pro Glu Arg Asn Thr Ser Ala Tyr Phe Asn
820 825 830

Ser Met Ile Gln Gly Tyr Thr Met Arg Lys Ser
835 840

<210> 20
<211> 839
<212> PRT
<213> Homo sapiens

<400> 20

Met Gly Pro Arg Ala Lys Thr Ile Cys Ser Leu Phe Phe Leu Leu Trp
1 5 10 15

Val Leu Ala Glu Pro Ala Glu Asn Ser Asp Phe Tyr Leu Pro Gly Asp
20 25 30

Tyr Leu Leu Gly Gly Leu Phe Ser Leu His Ala Asn Met Lys Gly Ile
35 40 45

Val His Leu Asn Phe Leu Gln Val Pro Met Cys Lys Glu Tyr Glu Val
50 55 60

Lys Val Ile Gly Tyr Asn Leu Met Gln Ala Met Arg Phe Ala Val Glu
65 70 75 80

Glu Ile Asn Asn Asp Ser Ser Leu Leu Pro Gly Val Leu Leu Gly Tyr
85 90 95

Glu Ile Val Asp Val Cys Tyr Ile Ser Asn Asn Val Gln Pro Val Leu
100 105 110

Tyr Phe Leu Ala His Glu Asp Asn Leu Leu Pro Ile Gln Glu Asp Tyr
115 120 125

Ser Asn Tyr Ile Ser Arg Val Val Ala Val Ile Gly Pro Asp Asn Ser
130 135 140

Glu Ser Val Met Thr Val Ala Asn Phe Leu Ser Leu Phe Leu Leu Pro
145 150 155 160

Gln Ile Thr Tyr Ser Ala Ile Ser Asp Glu Leu Arg Asp Lys Val Arg
165 170 175

```
Phe Pro Ala Leu Leu Arg Thr Thr Pro Ser Ala Asp His His Val Glu
            180                 185                 190

Ala Met Val Gln Leu Met Leu His Phe Arg Trp Asn Trp Ile Ile Val
        195                 200                 205

Leu Val Ser Ser Asp Thr Tyr Gly Arg Asp Asn Gly Gln Leu Leu Gly
    210                 215                 220

Glu Arg Val Ala Arg Arg Asp Ile Cys Ile Ala Phe Gln Glu Thr Leu
225                 230                 235                 240

Pro Thr Leu Gln Pro Asn Gln Asn Met Thr Ser Glu Glu Arg Gln Arg
                245                 250                 255

Leu Val Thr Ile Val Asp Lys Leu Gln Gln Ser Thr Ala Arg Val Val
            260                 265                 270

Val Val Phe Ser Pro Asp Leu Thr Leu Tyr His Phe Phe Asn Glu Val
        275                 280                 285

Leu Arg Gln Asn Phe Thr Gly Ala Val Trp Ile Ala Ser Glu Ser Trp
    290                 295                 300

Ala Ile Asp Pro Val Leu His Asn Leu Thr Glu Leu Gly His Leu Gly
305                 310                 315                 320

Thr Phe Leu Gly Ile Thr Ile Gln Ser Val Pro Ile Pro Gly Phe Ser
                325                 330                 335

Glu Phe Arg Glu Trp Gly Pro Gln Ala Gly Pro Pro Pro Leu Ser Arg
            340                 345                 350

Thr Ser Gln Ser Tyr Thr Cys Asn Gln Glu Cys Asp Asn Cys Leu Asn
        355                 360                 365

Ala Thr Leu Ser Phe Asn Thr Ile Leu Arg Leu Ser Gly Glu Arg Val
    370                 375                 380

Val Tyr Ser Val Tyr Ser Ala Val Tyr Ala Val Ala His Ala Leu His
385                 390                 395                 400

Ser Leu Leu Gly Cys Asp Lys Ser Thr Cys Thr Lys Arg Val Val Tyr
                405                 410                 415

Pro Trp Gln Leu Leu Glu Glu Ile Trp Lys Val Asn Phe Thr Leu Leu
            420                 425                 430

Asp His Gln Ile Phe Phe Asp Pro Gln Gly Asp Val Ala Leu His Leu
        435                 440                 445

Glu Ile Val Gln Trp Gln Trp Asp Arg Ser Gln Asn Pro Phe Gln Ser
```

450 455 460

Val Ala Ser Tyr Tyr Pro Leu Gln Arg Gln Leu Lys Asn Ile Gln Asp
465 470 475 480

Ile Ser Trp His Thr Val Asn Asn Thr Ile Pro Met Ser Met Cys Ser
485 490 495

Lys Arg Cys Gln Ser Gly Gln Lys Lys Lys Pro Val Gly Ile His Val
500 505 510

Cys Cys Phe Glu Cys Ile Asp Cys Leu Pro Gly Thr Phe Leu Asn His
515 520 525

Thr Glu Asp Glu Tyr Glu Cys Gln Ala Cys Pro Asn Asn Glu Trp Ser
530 535 540

Tyr Gln Ser Glu Thr Ser Cys Phe Lys Arg Gln Leu Val Phe Leu Glu
545 550 555 560

Trp His Glu Ala Pro Thr Ile Ala Val Ala Leu Leu Ala Ala Leu Gly
565 570 575

Phe Leu Ser Thr Leu Ala Ile Leu Val Ile Phe Trp Arg His Phe Gln
580 585 590

Thr Pro Ile Val Arg Ser Ala Gly Gly Pro Met Cys Phe Leu Met Leu
595 600 605

Thr Leu Leu Leu Val Ala Tyr Met Val Val Pro Val Tyr Val Gly Pro
610 615 620

Pro Lys Val Ser Thr Cys Leu Cys Arg Gln Ala Leu Phe Pro Leu Cys
625 630 635 640

Phe Thr Ile Cys Ile Ser Cys Ile Ala Val Arg Ser Phe Gln Ile Val
645 650 655

Cys Ala Phe Lys Met Ala Ser Arg Phe Pro Arg Ala Tyr Ser Tyr Trp
660 665 670

Val Arg Tyr Gln Gly Pro Tyr Val Ser Met Ala Phe Ile Thr Val Leu
675 680 685

Lys Met Val Ile Val Val Ile Gly Met Leu Ala Thr Gly Leu Ser Pro
690 695 700

Thr Thr Arg Thr Asp Pro Asp Asp Pro Lys Ile Thr Ile Val Ser Cys
705 710 715 720

Asn Pro Asn Tyr Arg Asn Ser Leu Leu Phe Asn Thr Ser Leu Asp Leu
725 730 735

Leu Leu Ser Val Val Gly Phe Ser Phe Ala Tyr Met Gly Lys Glu Leu
            740   745   750

Pro Thr Asn Tyr Asn Glu Ala Lys Phe Ile Thr Leu Ser Met Thr Phe
        755   760   765

Tyr Phe Thr Ser Ser Val Ser Leu Cys Thr Phe Met Ser Ala Tyr Ser
    770   775   780

Gly Val Leu Val Thr Ile Val Asp Leu Leu Val Thr Val Leu Asn Leu
785   790   795   800

Leu Ala Ile Ser Leu Gly Tyr Phe Gly Pro Lys Cys Tyr Met Ile Leu
                805   810   815

Phe Tyr Pro Glu Arg Asn Thr Pro Ala Tyr Phe Asn Ser Met Ile Gln
            820   825   830

Gly Tyr Thr Met Arg Arg Asp
        835

<210> 21
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 21
actttgagaa catgagtaat gacg   24

<210> 22
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 22
agtacccgga ctgcgtcgtc atta   24

<210> 23
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 23
cactagggtc atccttgctt tcag   24

<210> 24
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 24
agtcagggtg atgggcctga aagc   24

<210> 25
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 25
atgtggtgga ctggctgtac catc   24

<210> 26
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 26
ttgaagccct ccacgtgatg gtac   24

<210> 27
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 27
cacacggtga acaagatcac cttc   24

<210> 28
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 28
agtagcactg ctcggagaag gtga   24

<210> 29
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 29
atctaccaca tggacgagga ggag   24

<210> 30
<211> 24
<212> DNA
<213> Artificial Sequence

<220>

<223> Overgo probes

<400> 30
tgaccaggta cggcgtctcc tcct 24

<210> 31
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 31
agcgcgtcac gctggccgac ttca 24

<210> 32
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 32
ttgctgagca cgttcttgaa gtcg 24

<210> 33
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 33
cacgcctaca aattcttctt taag 24

<210> 34
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 34
agtcctggtc catggactta aaga 24

<210> 35
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 35
cttccactcc tgctgctacg actg 24

<210> 36
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 36
tgcctcgcag tccacgcagt cgta 24

<210> 37
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 37
aggtgcgccg cgtcaagggc ttcc 24

<210> 38
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 38
tcgtagcagc aggagtggaa gccc 24

<210> 39
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 39
gttcctggca tgggggagc cggc 24

<210> 40
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 40
gagcagcaca agcacagccg gctc 24

<210> 41
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 41
acagcccact agttcaggcc gcag 24

<210> 42
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 42
caggcccggg gtccccctgc ggcc 24

<210> 43
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 43
cccactggtt caggcctcgg gggg 24

<210> 44
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 44
aaagcaggcc aggggccccc ccga 24

<210> 45
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 45
aggcgctggt gcactgccgc acac 24

<210> 46
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 46
aagctgaccc aggagcgtgt gcgg 24

<210> 47
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 47
acagaggcac tggtgcactg ccgc 24

<210> 48
<211> 24
<212> DNA

<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 48
tgatccagga gtgcacgcgg cagt   24

<210> 49
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 49
accaatgcca cgctggcctt tctc   24

<210> 50
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 50
aagtgcccag gaagcagaga aagg   24

<210> 51
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 51
tggtacatgc tgccaatgcc acgc   24

<210> 52
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 52
aagcagagga aagccagcgt ggca   24

<210> 53
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgc probes

<400> 53
tacaaccgtg cccgtggcct cacc   24

<210> 54
<211> 24

<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 54
aggccagcat ggcgaaggtg aggc   24

<210> 55
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 55
tcatcacctg ggtctccttt gtgc   24

<210> 56
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 56
acattggcca ggaggggcac aaag   24

<210> 57
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 57
tgcagatggg tgccctcctg ctct   24

<210> 58
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Overgo probes

<400> 58
aggatgccca gcacacagag cagg   24